Intelligent ECG Processing and Abnormality Detection using Adaptive Ensemble Models

Diptangshu Pandit

A thesis submitted to the
University of Northumbria at Newcastle
for the degree of
Doctor of Philosophy

Department of Computer Science and Digital Technologies,
Faculty of Engineering and Environment,
Date : 14/04/2017

Acknowledgement

I would like to express my gratitude to my supervisors, especially to Dr Li Zhang, whose constructive guidance, support, and encouragement throughout the course were paramount to this work. Special thanks to Prof Samiran Chattopadhyay and Dr Chengyu Liu for helping me in some elements of the research.

I would also like to thank my parents for their constant support and encouragement throughout all my student life. Without their help, I would never have come this far. Special thanks to Melissa for the help with the thesis writing.

Declaration

I declare that the work contained in this thesis has not been submitted for any other award and that it is all my own work. I also confirm that this work fully acknowledges opinions, ideas, and contributions from the work of others.

Any ethical clearance for the research presented in this thesis has been approved. Approval has been sought and granted by the Faculty Ethics Committee on 02/2016.

I declare that the Word Count of this thesis is 40692 words.

Name: Diptangshu Pandit

Signature:

Abstract

This thesis explores the automated Electrocardiogram (ECG) signal analysis and the feasibility of using a set of computationally inexpensive algorithms to process raw ECG signals for abnormality detection. The work is divided into three main stages which serve towards the main aim of this research, i.e. the abnormality detection from single channel raw ECG signals.

In the first stage, a lightweight baseline correction algorithm is proposed along with a modified moving window average method for real-time noise reduction. Additionally, for further offline analysis, a wavelet transform and adaptive thresholding based method is proposed for noise reduction to improve signal-to-noise ratio.

In the second stage, a sliding window based lightweight algorithm is proposed for real-time heartbeat detection on the raw ECG signals. It includes max-min curve and dynamic (adaptive) threshold generation, and error correction. The thresholds are adapted automatically. Moreover, a sliding window based search strategy is also proposed for real-time feature extraction.

Subsequently, a hybrid classifier is proposed, which embeds multiple ensemble methods, for abnormality classification in the final stage. It works as a meta classifier which generates multiple instances of base models to improve the overall classification accuracy. The proposed hybrid classifier is superior in performance, however, it is dedicated to offline processing owing to high computational complexity. Especially, the proposed hybrid classifier is also further extended to conduct novel class detection (i.e. unknown newly appeared abnormality types). A modified firefly algorithm is also proposed for parameter optimization to further improve the performance for novel class detection.

The overall proposed system is evaluated using benchmark ECG databases to prove its efficiency. To illustrate the advantage of each key component, the proposed feature extraction, classification and optimization algorithms are compared with diverse state-of-the-art techniques. The empirical results indicate that the proposed algorithms show great superiority over existing methods.

Contents

List of Figures

List of Tables

Glossary of Associated Terms

Term	Description
ECG	Electrocardiography; A process of recording the electrical activity of the heart over a period of time. Recorded signal is known as ECG signal.
ECG Leads	A standard ECG contains 12 channels or Leads (Names: Lead I, Lead II, Lead III, aVR, aVL, aVF, V1, V2, V3, V4, V5, V6). (Boron & Boulpaep 2009) Each of which uses a pair of electrodes placed on human body to record ECG signal. Mostly Lead II signals were used in this thesis.
ECG Features	Features are valuable information which can be extracted from raw ECG signal. The extracted information is used further for abnormality detection.
ECG waves	ECG waves/sub-waves are the part of the ECG signals which can be uniquely identified and have some properties. Example: P, Q, R, S and T wave.
Classifier	A classifier is a supervised machine learning technique/model. It can identify to which of a set of categories a new observation belongs. A learning dataset and a training algorithm used while constructing the trained classifier.
QRS Complex	The QRS complex is a name for the combination of three (Q, R and, S waves) of the graphical deflections (sub-waves) seen on a typical electrocardiogram. It is usually the central and most visually obvious part in ECG signal generated during a single heartbeat.
R-Peak	R-Peak is the peak of R wave (sub-wave) in QRS complex. It generally represents the corresponding position of heartbeat in ECG signal.
RR Interval	The time interval between two consecutive R-Peaks

List of Publications

1. **A Lightweight QRS Detector For Single Lead ECG Signals Using A Max-Min Difference Algorithm**

   D Pandit, L Zhang, C Liu, S Chattopadhyay, N Aslam, CP Lim

   Computer Methods and Programs in Biomedicine, 2017

2. **Noise Reduction in ECG Signals Using Wavelet Transform and Dynamic Thresholding**

   D Pandit, L Zhang, C Liu, N Aslam, S Chattopadhyay, CP Lim

   Emerging Trends in Neuro Engineering and Neural Computation, 193-206, 2017

3. **Improved Abnormality Detection From Raw ECG Signals Using Feature Enhancement**

   D Pandit, S Chattopadhyay, L Zhang, N Aslam, C Liu

   Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), 2016

4. **An Efficient Abnormal Beat Detection Scheme From ECG Signals Using Neural Network And Ensemble Classifiers**

   D Pandit, L Zhang, N Aslam, C Liu, A Hossain, S Chattopadhyay

   Software, Knowledge, Information Management and Applications (SKIMA), 2014

Chapter 1 Introduction

This research investigates automated electrocardiogram (ECG) signal processing techniques and proposes intelligent ECG processing algorithms using evolutionary and ensemble methods. An ECG is the process of recording the electrical activities of the heart using electrodes that are placed onto the surface of the skin (Boron & Boulpaep 2009). It identifies and expresses the electrical impulses generated by the heart in a visual waveform which is used to help physicians to detect or diagnose various heart diseases. This research focuses on automated ECG signal processing on single lead ECG signals that are mainly used in mobile platforms.

The importance of automated ECG signal processing to detect heart abnormalities is discussed in this section along with the growing technologies in this area which may lead to more access to equipment for those at risk of such abnormalities. This chapter is further divided into four sections. Section 1.1 states the motivation behind this research and the areas where the outcome of this research can benefit. Section 1.2 describes the main goals of the research which involves improving overall ECG analysis in terms of accuracy and computational efficiency. In Section 1.3, research contributions are summarized. Section 1.4 outlines the thesis structure.

1.1 Motivation

Automated ECG processing allows more opportunities for people in the general population to access equipment which may assist in the early detection of cardiovascular diseases (CVD). Given the growing accessibility to mobile devices along with the growing technologies of mobile ECG sensors, it becomes easier for members of the public to utilise such equipment which can detect ECG abnormalities without visiting a medical doctor. This would increase the possibilities of early detection of diseases which currently claim more lives globally than any other diseases. There are many other motivations to this research; they are elaborated in the following.

1.1.1 Ever growing heart problems

Cardiovascular disease is a term used to refer to a class of disorders involving the heart and blood vessels. Cardiovascular diseases are the leading causes of death worldwide, resulting in around 17.3 million fatalities per year. This number is predicted to rise rapidly with around 23.6 million deaths per year expected by 2030 (Mozaffarian et al. 2015). Cardiovascular diseases claim more lives in total than those claimed by all forms of cancer combined. Operations and procedures offering medical intervention for the treatment of cardiovascular

diseases have increased by about 28% from 2000 to 2010, totalling 7.6 million in 2010. This demonstrates the growing attention being paid to these diseases. For example, each year in the U.S, approximately 735,000 people suffer from heart attacks. Of those, around 120,000 are fatal (Mozaffarian et al. 2015).

An additional strain is caused by the aging population. Due to healthier lifestyle options and the developments available in healthcare, people are living longer than they have in previous generations. This creates a pressure on the healthcare system to cater to the demands of elderly members of the public which involves many physical health issues (He et al. 2016). An aging population is contributing to the growing prevalence of CVD, due to the changes which occur in the heart as people age (WHO 2015). This concern is therefore a contemporary and increasingly important public health issue which must be managed. Mobile sensor devices could be used with this population to assist in managing their healthcare needs and preventing the development of untreated CVD.

1.1.2 Barriers to accessing healthcare

There are several issues creating increasing pressure on healthcare providers. For example, general practitioners in the UK have commented that their ever-growing workload is the biggest concern with regards to the delivery of healthcare (Hobbs et al. 2016). General practitioners are often the first point of contact for the general population to access healthcare. Therefore, due to increasing workload, it is more difficult to gain access to a doctor's appointment. Since general practitioners offer access to all other specialist services for patients, the struggles of gaining an appointment could result in progression of disease and those with CVD could be left undetected and untreated for longer than necessary. The employment of ECG analysis device which could detect these health issues could relieve the strain on both doctors and their patients, as those who use it could determine abnormalities before their visit to their doctor and therefore receive more efficient tailored care.

1.1.3 Availability of mobile ECG sensors

Wearable health monitoring is an emerging technology offering continuous monitoring of a patient's vital signs, including the ECG. This signal is widely adopted to identify heart abnormalities, and to diagnose and assess major health risks and chronic cardiac diseases. There are some restrictions for the wide use of deployed ECG monitoring systems. These include, imposed limitations on patients due to practicality, short battery life, lack of user acceptability and medical professional's feedback. Issues with patient confidentiality and the

privacy of essential data also offer a barrier to usage (Baig et al. 2013). Regardless of these limitations wearable health monitoring systems are becoming an emerging technology for continuous vital sign monitoring including ECG (Baig et al. 2013).

1.1.4 Availability and advancement of mobile devices

Increasingly, smartphones have become heavily ingrained into the daily life of the general population. Not only are the smartphones available getting smaller but also the computational power in a smartphone has grown exponentially; this growth is expected to continue (Tscherning 2011). This allows the opportunity for the integration of mobile sensors and the intelligent systems to be integrated in daily life. While a wide range of computing tasks can be accomplished by smartphones, some of the tasks require ever increasing computing power. These pose a challenge as the smartphones are resource constrained devices with limited computing power, memory, data storage and energy supply. However with the help of the cloud computing any type of computational work can be outsourced making virtually any task able to be accomplished through mobile device and internet connection (Khan et al. 2014).

1.1.5 The promotion of early detection of heart diseases

Early detection of heart problems has shown to contribute towards lower mortality rates in patients with these diseases (Kaplan et al. 1988). The use of ECG processing on mobile sensors offers greater accessibility to technology to detect heart abnormalities. Cardiovascular diseases are more prone to those whose lifestyle carries certain risk factors, such as smoking, physical inactivity, being overweight and drinking alcohol (WHO 2015). Therefore, those who may feel or may have been advised by a medical professional that their lifestyles carry a greater risk of developing cardiovascular diseases could seek out and utilise this equipment to monitor their hearts and seek medical advice if any abnormalities become apparent. This equipment allows the opportunity of continuous cardiac function assessment which in turn could lead to earlier implementation of medical treatment and overall reduced mortality rates.

The advancement in mobile technology along with the mobile ECG sensor opens the opportunity to aid healthcare by dealing with ever growing heart problems. Along with easily obtainable mobile ECG sensors, day-to-day-used mobile devices can be used to detect and diagnose cardiac problems in the early stages which might result in avoiding life threatening situations.

1.2 Objectives

The main goal of this research is to propose an intelligent decision support system for ECG signal processing and abnormality detection to improve upon existing mobile ECG detection processes. Developing complex algorithms can result in high computation complexity, which may lead to high battery usage and memory requirement on mobile devices. Mobile ECG processing is a complex process and can be divided into multiple stages. Each stage of the ECG processing involves corresponding algorithms or techniques. Starting from de-noising raw ECG signals to the detection of abnormal heartbeats, multiple techniques and novel algorithms have been proposed to improve performance in terms of R-peak and heartbeat abnormality detection without compromising computational cost. This vision results in an abnormal heartbeat detection scheme with superior performance. The objectives can be further elaborated as follows:

1.2.1 Abnormal heartbeat detection

Abnormal beat detection will involve multiple stages. To obtain higher accuracy, the raw signals need to be processed with sophisticated noise reduction techniques which focus on the signals obtained from mobile ECG signals rather than standard ECG electrodes. Efficient novel methods should be developed to extract the relevant parts of ECG signals. And finally, classification methods need to be used for robust abnormality detection.

1.2.2 Performance of the system

Making ECG processing light weight and real-time is another goal of this research. This is to make sure the proposed system can be deployed to diverse processing platforms which might have limited resources (e.g. processing power and memory). This will also address the barrier of the equipment causing practical limitations to the wearer.

1.2.3 Online and offline processing

The goal of making an online detection system is to establish real-time heartbeat abnormality detection. This can help with life threatening situations by creating a notification which is accessed by the emergency services. However, we also aim to realize an offline system for thorough processing of the obtained signals. The offline analysis will further help determine and analyse the specific type of abnormality in the ECG signal as opposed to the presence or absence of an abnormality.

The proposed research aims to address the requirement of both online and offline ECG signal processing and the detection of abnormal heartbeats. For instance, the online ECG processing may require the system to be as simple as possible in terms of computational complexity and memory consumption, while the offline process may require sophisticated methods to conduct thorough processing of the signals to achieve higher accuracy in abnormality detection. This research will take both aspects into account and address the best trade-off between the system performance and computational efficiency.

1.2.4 Unknown beat detection

Because of the diversity of heartbeat abnormalities, it could be highly possible that the training dataset may only include partial types of abnormal heartbeats. This may leave the remaining types of heartbeats unknown to the system, but could appear later. It is critical for the system to correctly judge whether a test heartbeat belongs to a known category included in the training set or a newly arrived novel class type. Therefore, the final objective of this research is to identify the arrival of unknown heartbeats automatically. As a general novel class detector, it should work on other diverse databases for novel class detection.

1.3 Contributions

In this research, multiple algorithms have been proposed with the intention of contributing to the improvement of automated ECG processing and classification. The research has focused on the processing of Lead II ECG signals and proposed algorithms have been tested by using several standard databases. The main contributions have encompassed multiple domains, including, signal processing, feature extraction, classification, and optimization.

1.3.1 ECG Noise reduction

An online baseline removal and noise reduction algorithm is proposed to deal with the various noises which may be present while acquiring ECG signals using mobile ECG sensors. The noise reduction algorithm is used as a prerequisite of QRS detection algorithm for further analysis as any noise can result in performance decrease in subsequent ECG processing stages. This online algorithm is developed specially for mobile ECG sensors which only include certain types of noise.

Another pre-processing algorithm is also proposed which aims to increase the signal to noise ratio of noisy ECG signals. In achieving this target, the ability of efficiently reading ECG signals could be improved. The algorithm is developed using wavelet transform and dynamic

thresholding approaches. The proposed algorithm has yielded impressive results in reducing the noise present in signals. This process is also required when dealing with offline signals or pre-recorded signals to improve its quality and readability.

1.3.2 Light-weight QRS Detection

The next step in ECG processing after noise reduction is to identity the R-peak from the raw ECG signals. We developed a light weight QRS complex (i.e. heartbeat occurrence region) detection algorithm, which utilizes a max-min difference curve generation technique to detect R-peaks from Lead II ECG signals. This algorithm can be used online as it uses a sliding window based approach. It is also very convenient to be used on mobile platforms while processing single lead ECG signals captured using mobile ECG sensors. The algorithm has been tested on several standard ECG databases for R-peak detection and has also been compared with some of the latest methods in terms of accuracy and computational complexity to prove superiority and usefulness of the proposed method. This method can be used directly as the heartbeat variability analysis or as a prerequisite of feature extraction process from raw ECG signals.

1.3.3 Feature Extraction and Enhancement

A simple feature extraction algorithm is also proposed to locate the P, Q, R, S, and T waves within Lead II ECG signals. The feature extraction algorithm uses a window based search technique around the QRS location. As the proposed method uses a local window search approach and a small buffer of signals, it can be easily implemented on and is most suitable for mobile devices. The algorithm is used on multiple databases for feature extraction and the extracted features are tested on multiple databases resulting in high abnormality detection rates.

Subsequently, a feature enhancement (improvement) technique is introduced that adds more medically valuable information along with the standard (P, Q, R, S and T) wave locations. The proposed technique includes the overall energy in each sub-wave which is used to improve abnormality detection accuracy.

1.3.4 Hybrid Classifier

A hybrid classifier is proposed which can be used for abnormality detection along with the extracted ECG features. The proposed classifier uses multiple base ensemble classifiers, each of which extracts a different scenario of the original training dataset. Finally, collaboratively the trained base ensemble classifiers are used to determine the final class label by using a

weighted majority voting method. The weights of the proposed base classifiers are optimized by an evolutionary optimization algorithm. The proposed algorithm works as meta classifier which can be used along with any base classifier such as decision tree, artificial neural network, support vector machine etc. In our experiments, we implemented the meta classifiers using decision tree learning techniques and the results show that the proposed algorithm shows higher accuracy than the standard single and ensemble classification methods. The proposed hybrid classifier has also been evaluated with other UCI benchmark datasets (Bache & Lichman 2013) to indicate its efficiency.

1.3.5 Novel Class Detection

A novel class detection algorithm is proposed to detect unknown type heartbeats as an extension to the abovementioned hybrid classifier. The algorithm is constructed using clustering, and the hybrid classifier. A novel evolutionary algorithm is proposed for base classifier weight and novelty parameter optimization (see Section 1.3.6). To demonstrate the efficiency of this method, the algorithm is tested against multiple databases, including the derived ECG dataset and other diverse UCI datasets.

1.3.6 Repulsive Firefly

A new improved firefly algorithm (FA) is proposed to conduct optimization of classifier parameters. This can be used as a general purpose and standalone optimization algorithm. The proposed algorithm utilizes repulsive force and scattering behaviours along with the original FA mechanism to solve diverse optimization problems. Besides the original attractiveness movement behaviour of the FA, the proposed optimization algorithm utilizes a repulsive force strategy (known as RFA) to reach global optimality in fewer iterations and scattering mechanisms (known as SRFA) to overcome premature convergence. Eleven different benchmark optimization functions are used for the evaluation of the proposed RFA and SRFA. The empirical results indicate that both proposed algorithms outperform the original FA and other conventional and state-of-the-art optimization algorithms such as, simulated annealing, particle swarm, bat swarm, cuckoo search, dragonfly, and ant-lion optimization, significantly. SRFA integrated with both repulsive and scattering behaviours also outperforms RFA for the evaluation of nearly all of the test benchmark functions.

1.4 Thesis outline

Chapter 2 focuses on the introduction of automated ECG processing. The introduction of an ECG signal is provided along with its components. Furthermore, the automated ECG analysis is described step by step (including e.g. pre-processing, feature extraction etc.). The common methods used in each corresponding stage are also explored.

In Chapter 3, we discuss the current research on ECG processing in the field. The relevant research has been divided into several categories according to the key steps of the ECG analysis. The chapter concludes with current challenges in ECG processing and viable solutions.

Chapter 4 describes the first key step of this proposed research, i.e. the ECG pre-processing method along with experimental results. A noise reduction method is proposed using wavelet transform and dynamic thresholding for offline processing. Additionally, a baseline removal and smoothing technique is also developed for preparing the raw ECG signals for online processing.

In Chapter 5, multiple ECG feature extraction methods are proposed and experimented upon. First, a light weight sliding window based QRS detection is developed for Lead II ECG signals, which utilizes max-min difference curve generation. Next, two feature extraction methods are introduced. The first method is developed to identify all of the features (baseline of the signal and the locations of P, Q, R, S, T waves) from raw signals while the other algorithm uses the R (wave) peak location as the base input to locate other signal peaks (such as P, Q, S and T peaks). Both the algorithms employ the search window around QRS locations to find corresponding features in the ECG signals.

In Chapter 6, two experiments have been conducted to show how generating features can improve overall disease detection accuracy. The first experiment employs all of the standard features (e.g. relative locations and amplitudes of P, Q, R, S and T waves) whereas the second experiment involves additional features (e.g. overall energy in each of the sub-waves) to improve the detection results.

Chapter 7 discusses ECG classification and novel class detection. A hybrid ensemble classifier is proposed to improve the classification performance of single model classification techniques. This classifier is not only capable of increasing classification accuracy in comparison to other standard ensemble methods, but it is also able to detect the arrival of novel abnormality classes which may not be present in the training database.

The optimization algorithms (i.e. RFA and SRFA) embedded in the novel class detector are discussed in Chapter 8. To prove efficiency of the proposed evolutionary algorithms, multiple test benchmark functions have been used for evaluation. The results of multiple optimization algorithms are also compared with those from other classical search methods and several state-of-the-art FA variants.

Finally, Chapter 9 summarizes the thesis along with research contributions. It also provides indications for the possible future directions of this research.

Chapter 2 Background

An ECG is the recording of the electrical activities of the heart. The signal expresses the electrical impulses generated by the heart in a visual waveform which is used by physicians to monitor the myocardial activities of the patient (Boron & Boulpaep 2009). This helps the physicians to diagnose various heart diseases if present. ECG signals can have multiple channels/leads. In a conventional 12-lead ECG, 10 electrodes are placed on the surface (limbs and chest) of a human body non-invasively to record the ECG signals while mobile ECG sensors generally only provide one or two leads.

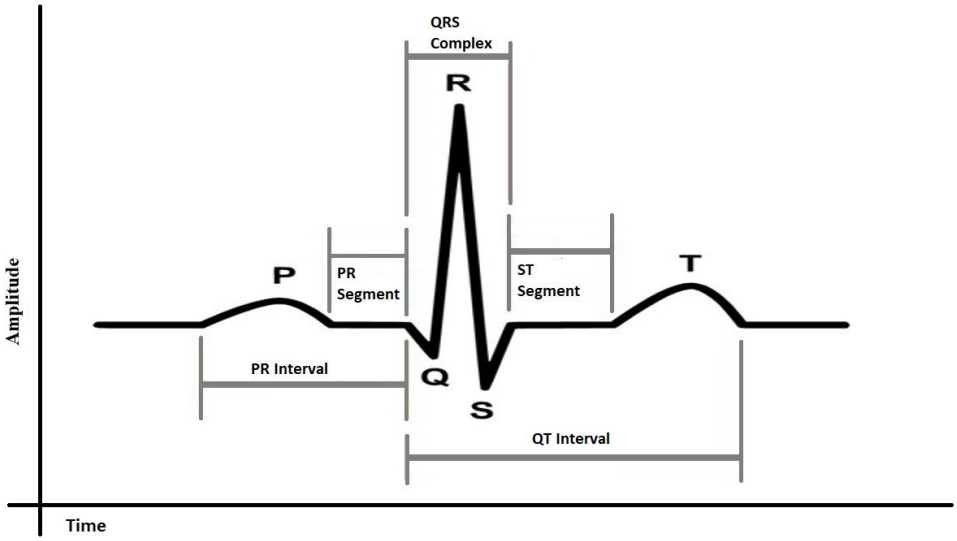


Figure 1 : Ideal Lead II ECG waveform.

Traditional ECG machines use twelve leads to record 360-degree heart activity. This provides detailed electrical activities of heart from multiple directions (Boron & Boulpaep 2009). However, mobile sensors most commonly use only 1 lead (generally lead I or lead II). This makes the processing of the ECG more difficult as it conveys significantly less medical information than the 12 lead signals which makes it harder for both the physicians and automated algorithms to diagnose any possible abnormality present in the signal.

An ideal one-channel (Lead II) ECG consists of five regions; namely P, Q, R, S and T waves (Boron & Boulpaep 2009) as shown in Figure 1. Meaningful information can be extracted from raw ECG signals by finding these wave positions, durations, amplitudes, and shapes. Furthermore, accurate detection of waveforms and processing of these signals are challenging

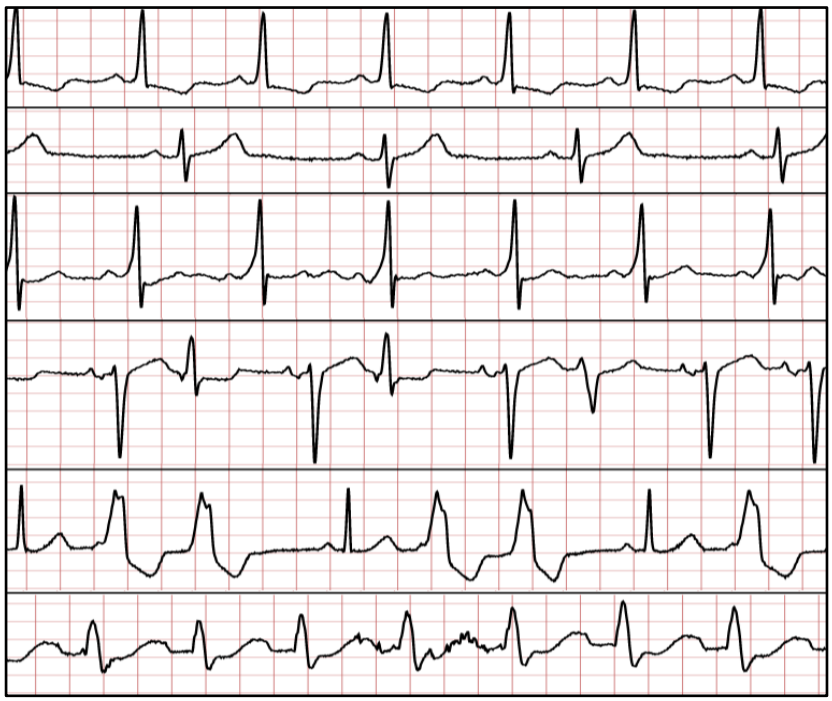


Figure 2 : Real ECG samples taken from database

tasks as the natural raw signals taken from a human body are more diverse and sometimes differ significantly from the ideal case as shown in Figure 2. The plotted graphs are real ECG signal plotted from MIT-BIH Arrhythmia database (Moody & Mark 2001) consisting of normal heartbeats (top three) and abnormal heartbeats (bottom three). From the information in Figure 2, it can be seen that even the normal heartbeats (in ECG signal) differ greatly from each other.

The raw signal should go through multiple stages before any abnormality can be detected automatically without the need for human intervention. The main stages are 1. Signal pre-processing, 2. Feature extraction, 3. Feature enhancement or reconstruction and 4. Classification. In the first stage, all unwanted noise components present are cancelled and the important signal components are isolated. Next, in the feature extraction stage, ECG features (e.g. wave locations, amplitudes, shapes etc.) are extracted from the noise free (or less noisy) ECG signal. Feature enhancement is an optional stage which helps with enhancing desired components of related features according to the need of the subsequent algorithm. Derived features are generated and optimized to be used in the next stage. Finally, the extracted features are used for classification and abnormality detection. Single heartbeats or a part of the signal can be classified for abnormality detection. A flowchart of the entire process is shown in Figure 3 and each stage is discussed in detail below.

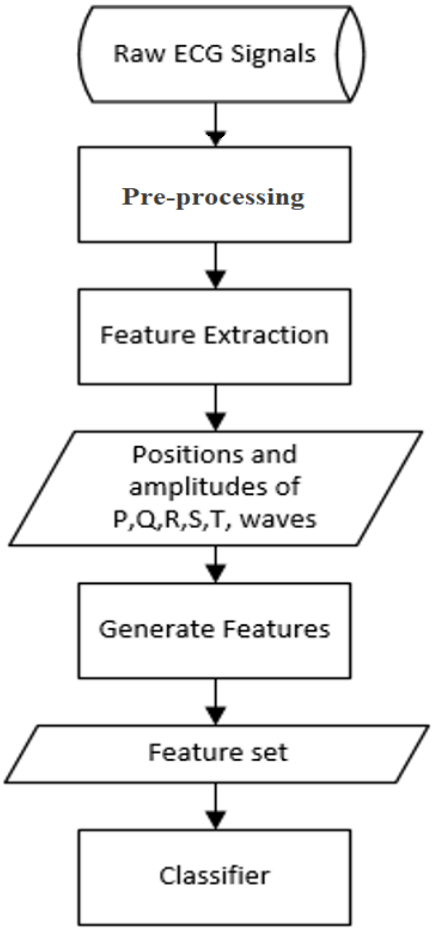


Figure 3 : ECG processing flowchart

2.1 Pre-processing

The first stage of ECG processing is denoising the signal. The noise removal process separates the valid and necessary parts of the signal by cancelling all kinds of unwanted artefacts which have not been influenced by electric activities of the heart. The most common noise categories include, Powerline Interference, Motion Artefacts, Baseline Wander, Muscle Contractions, and Electrode Contact Noise. They are explained in detail in the following.

2.1.1 Powerline Interference

This type of noise consists of 50Hz or 60Hz (and multiples) pickups and harmonics. The amplitude of the noise varies up to 50% of the ECG amplitude peak (Thalkar 2013). The common causes include improper grounding, interference with other electrical equipment (e.g. motors, air conditioners), and stray effect of the alternating current etc. An example signal is shown in Figure 4(a). Also all the plotted examples in Figure 4 are taken from MIT-BIH Arrhythmia database (Moody & Mark 2001).

2.1.2 Motion Artefacts

The transient baseline changes in ECG signals due to the difference in the electrode-skin impedance with electrode motion, are known as motion artefacts (Friesen et al. 1990). This type of noise might cause very large spikes. This has a frequency typically less than 0.5 Hz except for the abrupt shifts (due to motions). A signal with motion artefact present is illustrated in Figure 4(b).

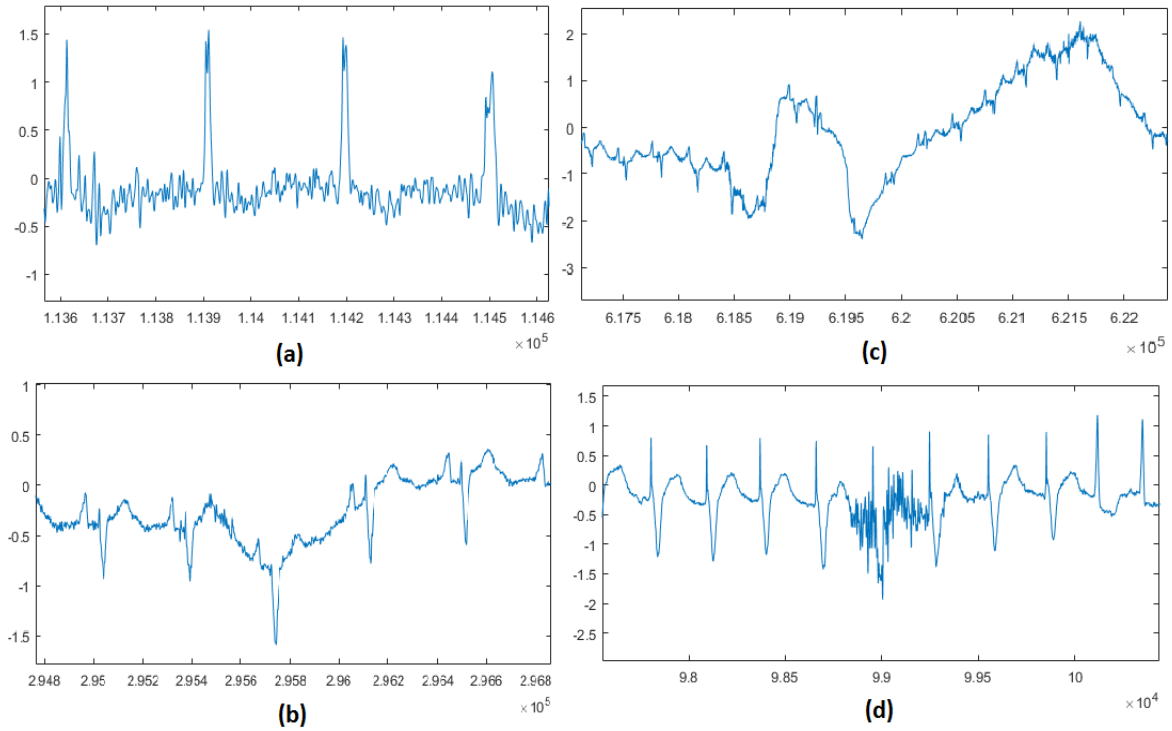


Figure 4 : Different types of noises in ECG signals.

(a) Powerline Interference, (b) Motion Artefact, (c), Baseline wander and (d) EMG noise. y-axis represents the amplitude in mV and x-axis denotes the sample number.

2.1.3 Baseline Wander

Respiration or body movement causes the Baseline Wander of low frequency noise. Baseline Wander causes problems for detecting R-peaks of the signal. Sometimes, because of this Baseline Wander, the T peak might appear higher than the R-peak (Kaur et al. 2011). Figure 4(c) shows an example of an ECG signal with Baseline Wander.

2.1.4 EMG Noise

Muscle contractions noise, also called Electromyogram (EMG) noise, is caused by the electrical activities of body muscles (Raphisak et al. 2004). This type of noise might have a frequency from 20 to 10KHz. The amplitude of noise can go up to 10% of the QRS amplitude. This type of noise is shown in Figure 4(d).

2.1.5 Electrode Contact Noise

This category of noise appears due to discontinuous connection between electrode and patient skin contact during ECG acquisition. Electrode contact noise results in discontinuous ECG signals and variation wave peaks.

Overall, the noise removal is not directly related to the feature extraction process; however, it impacts highly on feature extraction and classification of ECG signals when a considerable amount of noise exists in the corresponding raw ECG signal.

2.2 Feature Extraction

In this phase, firstly QRS complexes are detected from the raw ECG signal which determines the location of heartbeats in the signal. In the next stage, the medically prominent features are extracted as marked in Figure 1. These features include durations of P, T waves, PR interval, PR segment, QRS complex, ST segment, QT interval etc. Each of the important features will be discussed further in the following.

2.2.1 P wave

The P wave is the result of depolarization of the atria which spreads from the SA to AV node, and from the right to the left atrium.

The P wave is generally upright in Lead II; an inverted P wave axis might be a result of an ectopic atrial pacemaker. A P wave with a long duration may represent atrial enlargement. A tall P wave is caused by large right atrium while a notched bifid is resulted by a large left atrium.

2.2.2 PR interval

The PR interval is measured by calculating the distance from the beginning of the P wave to the beginning of the QRS complex. The time which the electrical impulse takes to travel from the sinus node through the AV node is represented by this interval.

A shorter PR interval (less than 120ms) means that the electrical impulse is bypassing the AV node which could be a sign of Wolf-Parkinson-White syndrome. A PR interval which is consistently longer than 200ms might indicate first degree atrioventricular block. The PR segment, which is measured as the portion of the tracing after the P wave until the QRS complex, is generally completely flat in normal conditions. However, it may be depressed in pericarditis.

2.2.3 QRS complex

The QRS complex is generated as a result of the rapid depolarization of the right and left ventricles. The QRS complex generally has a much greater amplitude than the P-wave, since the ventricles have a large muscle mass compared to the atria.

If the QRS complex appears wider than normal (longer than 120ms) then this suggests disruption of the heart's conduction system, such as in left bundle branch block, right bundle branch block, or ventricular rhythms such as ventricular tachycardia. Longer QRS complex can be present because of metabolic issues such as severe hyperkalemia, or tricyclic antidepressant overdose. Left ventricular hypertrophy could cause an unusually tall QRS complex while a pericardial effusion or infiltrative myocardial disease could result in a very low-amplitude QRS.

2.2.4 ST segment

The ST segment is the portion of an ECG signal that connects the QRS complex and the T wave; it signifies the period of ventricular depolarization.

It is typically isoelectric, however, could be depressed or raised because of myocardial infarction or ischemia. ST depression can also be the result of left ventricular hypertrophy or digoxin. Raised ST can also be caused by pericarditis, Brugada syndrome, or can be just a normal variant.

2.2.5 T wave

The repolarization of the ventricles is represented by the T wave. It is normally upright in all leads except aVR (augmenter Vector Right) and lead V1 (chest lead). Reversed T waves can be a mark of myocardial ischemia, left ventricular hypertrophy, high intracranial pressure, or metabolic abnormalities. Hyperkalemia or very early myocardial infarction can cause parked T waves.

2.2.6 QT interval

The QT interval is calculated as the distance from the start of the QRS complex to the end of the T wave. Suitable ranges vary with heart rate, because of which it must be modified accordingly by dividing by the square root of the RR interval.

A lengthy QT interval indicates a risk factor for ventricular tachyarrhythmias and sudden death. Extended QT could be caused because of genetic syndrome, or as a side effect of certain medicines. An oddly short QT can be seen in severe hypercalcemia.

In the feature extraction stage, all these features (medically valuable information) are extracted for further processing. The final abnormality detection results of the entire process are strongly dependent on this stage. A good feature extraction algorithm can result in a highly accurate ECG diagnosis system. Although, most of the times no feature extraction algorithm can be as good as a human physician because of variance in real ECG waveforms for both normal and abnormal heartbeats.

2.3 Feature Enhancement

Feature enhancement or feature reconstruction is an optional step which is only used for improving accuracy in the later stages. Standard features are medically important, however, finding accurate measures are sometimes computationally infeasible or very expensive. Some alternative features can be extracted which convey a subset of the medically valuable information. Extra features can be generated to further improve the outcome in terms of accuracy and robustness. For example, finding the exact boundaries of P or T wave is difficult. Especially, T wave boundaries are very hard to be determined correctly as the T wave comes in multiple shapes. A large amount of other information, such as the shape of the waves, presence of notch in R wave etc. can be derived/generated after the feature extraction process, which will be useful for subsequent abnormality detection from ECG signals.

2.4 Classification

Finally, extracted and enhanced features are used to determine whether a particular heartbeat or a series of heartbeats is normal or abnormal. If there is any abnormality present, the next step will be to try to identify the type of the abnormality. The simple steps are illustrated in Figure 5. The main reason of the two-step classification as indicated in Figure 5 is owing to prioritisation. For instance, it is more important to find out whether a heartbeat is normal or

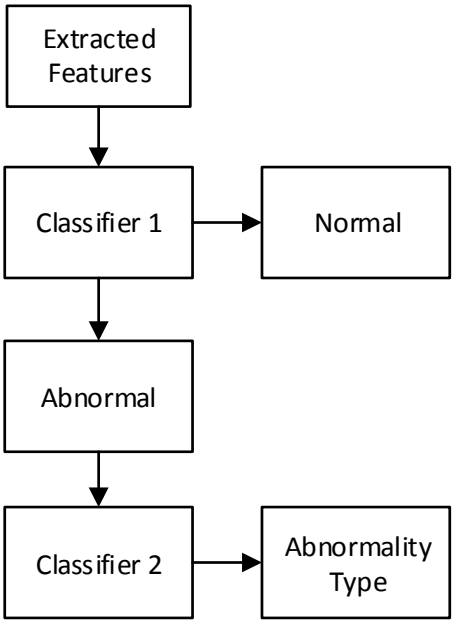


Figure 5 : Diagram of classification.

abnormal than to find the type of abnormality. If the heartbeat is abnormal then the patient can access a doctor's attention regardless the type of abnormality.

Multiple types of classification algorithms can be applied collaboratively to determine the presence of abnormalities or to precisely find out the abnormality type. Before performing classification tasks, classifiers need to be trained with the use of existing manually annotated ECG databases. After all the classifiers have been successfully trained, they can then be used for future classification. Importantly, some of the classifiers can even predict the presence of completely new types of abnormality in the heartbeat which are currently absent from the training dataset.

2.4.1 Single Classifier

There are many standard classifiers which can be used for this purpose, for example, Artificial Neural Network, Support Vector Machine, Decision Tree, Random Forest, Naive Bayes etc. These classifiers are computationally efficient, however, using single classifier is not always the best choice when there is a demanding request on the performance of disease detection.

2.4.2 Ensemble Classifiers

Ensemble classifiers can be used to gain more accurate classification performance of the ECG signals. Ensemble models use multiple base classifiers and their combined knowledge to

determine the final classification result. Some of the widely-used ensembles include Bagging (Bootstrap aggregating), Boosting, Stacking etc.

2.4.3 Novel Class Detection

ECG signals are very dissimilar in nature. For example, even normal ECG waveforms vary from person to person. Training the classifiers with an existing set of ECG data may not be sufficient for the classifier to successfully determine the correct output classes. It is also difficult to implement a classifier which can correctly identify all the types of abnormalities that might be present in ECG signals. In these cases, novel class detection can provide an appropriate solution. A novel class detector identifies the test cases that do not match the current class definitions available in the training database. These new cases would be most appropriate to be automatically identified to reduce error.

2.5 Conclusion

Automated ECG processing is a multi-stage process, which mainly consists of noise reduction, feature extraction and improvement (feature generation or feature selection) and classification. In each stage of ECG processing different methods are used. For example, in the noise reduction stage, one or multiple filtering techniques are used to cancel out the noise. In the feature extraction stage, various methods are used to extract medically valuable information (or features). And finally, these features are used for classification to find out any abnormalities that might be present in the ECG signal.

The overall process of automated ECG overlaps on multiple domains including signal processing, classification techniques, and soft computing domains. As the main goal of this research is to make automated ECG analysis more suitable for diverse platforms, multiple algorithms are proposed in the related domains to improve efficiency by either reducing resource (computational) usage or by improving overall accuracy of the proposed algorithms. Related work on each stage of ECG processing is discussed in detail in the following chapter.

Chapter 3 Related Work

Automated ECG processing is a popular and well-explored area in literature. In this chapter, the state-of-the-arts of mobile ECG systems are first discussed and then, the related work for each stage of the automated ECG processing is presented. The chapter is divided into multiple sections with the first section discussing various available mobile ECG systems and the rest of the sections dedicated to the advancements on each stage of ECG processing.

3.1 Mobile ECG processing systems

Various mobile ECG analysis systems have been developed over the past two decades for disease diagnosis. A comprehensive survey of wearable and wireless ECG monitoring systems can be found in the work of Baig et al. (2013). Furthermore, multiple aspects are also discussed including the efficiency of the systems, user acceptability, strategies, and recommendations for improvement over current systems. The wearable monitoring systems include bio-medical sensors (ECG sensors), communication techniques (e.g. Bluetooth, IEEE 802.11 etc.), complex computing (e.g. phone, computer) and artificial intelligence methods to produce the intended results.

There are several review works regarding wearable monitoring systems in smart homes (Chan et al. 2008), wireless body area networks (Ullah et al. 2012), wireless sensor networks for healthcare (Alemdar & Ersoy 2010) and mobile telemedicine (Lin 2012). A few more reviews of ECG signal collection, processing, used technologies, limitations and future improvement recommendations from 2000 to 2007 are also available (Addison 2005; Davenport et al. 2006; Poli et al. 2003). However most of these techniques are outdated because of the advancement in multiple corresponding underlying technologies in recent years.

3.1.1 Wearable Systems

ECG data acquisition can be performed through various methods. Textile sensors are popular in this domain. For example, ECG has been collected through smart shirts (Lee & Chung 2009; Bianchi et al. 2010). In this type of system, ECG sensors are generally placed inside the garment touching the human skin in corresponding places while the connection can be done with or without wires. ‘LOBIN’ is an ECG monitoring smart shirt which utilizes combination of e-textile sensors and wireless sensor network (López et al. 2010). The mobile ECG system has also been adapted for use over a longer period of time continuously for monitoring purpose with the help of planar-fashionable circuit board technology (Yoo, Yan, Lee, Kim & Yoo

2009). A wearable system, known as a mobile personal trainer ('MOPET'), is developed for supervising physical fitness activities which can help in better physical training and motivating users (Buttussi & Chittaro 2008). A T-Shirt based system is developed which could be used for capturing ECG signal as well as for measuring respiratory rate (Jourand et al. 2009). A novel hand-held device, 'Blue Box' (Pollonini et al. 2012), can collect and wirelessly transmit key parameters of ECG (features), bio-impedance and photoplethysmography. It can also compute many features including QRS duration and RR interval, heart rate (HR) and systolic time intervals. Additionally, it can assess their values based on the cardiac output computed by echo-doppler. The work of Yoon et al. (2011) uses a novel way by finding optimal condition difference between three textile electrodes and skin to measure the degree of skin hydration. A commercially accessible biofeedback gadget which aimed to improve heart rate fluctuation, called 'StressEraser ®', was proposed by Heilman et al. (2008). The 'E-chair' was proposed by Chang et al. (2013) to record physiological signals including body temperature, blood pressure, heart rate, body fat percentage and some other properties by installing multiple data measurement modules (sensors) into a chair. The proposed system can also display the computed results and diagnose simple problems in real time (Chang et al. 2013). Similarly, another system was developed for wireless health monitoring using wearable belt (Sardini & Serpelloni 2010). The monitoring system is very useful for the patients who are staying at home. A wearable physiological monitoring system, smart vest (Pandian et al. 2008), is developed which uses a number of sensors integrated into cloth fabric. The developed system is simultaneously capable of collecting multiple bio-signals in a non-invasive and unobtrusive way. Additionally, it can record ECG signals without using a gel substance. The recorded ECG is free from baseline noise and motion artefacts because of inbuilt hardware-filters (high pass, low pass, and notch filters). 'BIOTEX' (Coyle et al. 2010) is a bio-sensing textile sensor which can measure physiological parameters along with chemical composition of body fluids (e.g. sweat). 'BIOTEX' consists of sensors for detecting conductivity, sodium, and pH. 'MEMSWear' (Tay et al. 2009) uses a similar approach for remote monitoring of vital bio signs. It combines a wireless body area network and a personal digital assistant for this purpose.

3.1.2 Wireless and Mobile Systems

Mobile ECG analysis systems can be implemented using several wireless technologies. Wireless body area network (WBAN) is a popular choice in this area. WBAN enables the integration of multiple elements for monitoring physiological activities including miniaturized components, low-power consuming sensors and intelligent systems. ECG data collection has

been successfully practiced using contactless (Yoon et al. 2009) and leadless ECG monitoring (Scherr et al. 2008). Disadvantages and possible adverse effects of electrodes on the human body are also addressed (Winterhalter et al. 2008) with recommendations for the best electrode positioning (Puurtinen et al. 2009). In acknowledgement of obstructive sleep apnea, a cost effective, real-time monitoring system, 'MedAssist' has been developed (Bsoul et al. 2011). A energy-efficient low-complexity ECG compression has been proposed for signal acquisition and compression using shimmer WSBN mote (Mamaghanian et al. 2011). A portable ECG monitoring framework has been developed using a Holter-based ECG monitoring system as well as two smartphones for cardiovascular diagnosis (Oresko et al. 2010). It is able to conduct real-time and continuous monitoring to record patients' ECG signals. Accuracy and power usage problems have been resolved to some extent using power-saving, lightweight, and RFID-based wireless or USB ECG devices (Lin et al. 2010), however, there is a possible threat of data security (Chung-Chih et al. 2008).

There are several mobile device-based ECG feature detection systems. Wireless body area network is widely used in most of such devices. Some of them utilize wavelet transform (Kim et al. 2011). The examples include an image-guided ECG signal detection (Barnwell et al. 2012), a BSN-based context aware QRS detection (Kim et al. 2011), and e-technology in a unified DICOM format (Hsieh & Lo 2010). These systems use completely mobile based technology. A multi-parameter remote monitoring system has been developed for physiological signals based on the browser/server model. The parts of the system include a server monitoring centre, internet network and PC-based multi-parameter monitors on the world-wide-web which uses multimedia messaging service (MMS), general packet radio service (GPRS) and global positioning system (GPS) for transmission of acquired ECG data and storing in the Holter monitor using the Internet. A medical-embedded device for individualised care ('MEDIC') was developed, which uses an innovative software architecture to enable sensor management along with disease prediction using personal digital assistant (PDA) or a cell phone (Wu et al. 2008). A mobile medical device, 'HeartSaver', was developed (Sankari & Adeli 2011). It monitors ECG signals in real time and automatically detects several cardiac problems. It is a mobile-based (Android) application software which sends a text alert containing the patient's condition and location to a doctor when required. For continuous ECG monitoring system, an attachable ECG sensor adhesive bandage was implemented which uses Planar-Fashionable Circuit Board technology. It integrates a low cost sensor chip that is bonded

on fabric. Data is gathered wirelessly using dry electrodes for less skin irritation and long-term monitoring (Yoo et al. 2009).

3.1.3 Discussions

There are a variety of problems in current mobile ECG processing systems. For example, one of the common problems of all systems which use commercial wet ECG electrodes is that they dry out soon and sometimes the patient's skin is affected. Wireless ECG systems also suffer from high battery consumption because of power depletion due to wireless transmission. Similar problems can arise in battery operated mobile devices, which consume lot of power, while processing the ECG signals. This research is focused on making cost-effective ECG processing methods which will produce highly accurate results without compromising much of the resource usage. In the next section, currently used ECG analysis methods are discussed. The methods are grouped together based on each key stage of ECG analysis.

3.2 Mobile ECG processing algorithms

ECG processing algorithms play a vital role underneath all the ECG processing systems to detect abnormalities. Even for the ECG recoding and monitoring with human intervention some pre-processing of the signal is needed for noise reduction. Advanced challenges lie in feature extraction and intelligent disease diagnosis (abnormality detection) without any human intervention.

3.2.1 ECG pre-processing

As mentioned earlier, noise removal is a very important step and impacts highly on the subsequent processing of the ECG signals. It has drawn rising research interests in the related field. Especially noise reduction using filtering has become a very popular research area in signal processing domain.

Filters can be divided into two categories based on impulse response types, which are discussed below along with their advantages and limitations.

3.2.1.1 FIR Filters

In signal processing, a Finite Impulse Response (FIR) filter is a filter which produces finite duration impulse as a response to any finite length input.

An FIR filter has multiple properties that may make it more preferable than IIR filters. For example, it does not require any feedback. This means that rounding errors do not reflect in the

final result because of compound summed iterations (like in IIR). In each calculation, the same relative error occurs. This also makes the filter easier to be implemented. FIR can be easily designed to be linear phase by making the coefficient sequence symmetric. However, there are some disadvantages too. One of the disadvantages of the FIR filter is that it requires considerably more computational resources compared to those of an IIR filter.

3.2.1.2 IIR Filters

In signal processing, an Infinite Impulse Response (IIR) filter is a filter which produces infinite duration impulse as a response. Analog electronic filters, for example, composed of resistors, capacitors, and/or inductors are generally IIR filters.

The main advantage of IIR filters over FIR filters is that they are easy to implement. If IIR is implemented in a signal processor, correspondingly fewer calculations per time step will be needed compared to FIR version. Using IIR filters is beneficial where saving computational resource is a huge factor.

There are a few disadvantages as well. When it is required to design a particular frequency response unlike the usual cases (high-pass, low-pass, notch, etc.), it is harder to implement in IIR than in FIR. Another problem is the potential for limit cycle behaviour when idle, due to the feedback system in conjunction with quantization.

Different types of filters have been used for pre-processing of the ECG signals such as band-pass filters, notch filter, FIR and IIR filters etc. A comprehensive study of employed notch filters is presented by Bai et al. (2004). In their work, mean square error was used in performance comparison while demonstrating the adverse effect of higher order filter. Rani et al. (2011) presented a comparison of multiple FIR and IIR filters and illustrated the superiority of IIR filters over FIR filters in terms of resource requirements. Other types of filters include adaptive and morphological filters (An-dong et al. 2012; Tadejko & Rakowski 2007), median and band pass filters (Pan & Tompkins 1985; Tamil et al. 2008), and quantization (Chouhan & Mehta 2008). Recently, wavelet transform has gained popularity in automated ECG processing. As ECG is a non-stationary signal, wavelet methods are highly suitable for automatic ECG signal processing. Many wavelet noise removal techniques are also proposed in literature (Kadambe et al. 1999; Kabir & Shahnaz 2012; Garg et al. 2011; Khan et al. 2011; Kaur et al. 2011; Li et al. 2012; Thalkar 2013; Tiwari & Dubey 2014).

3.2.2 ECG feature extraction

Computerised ECG analysis has been widely studied resulting in many advanced methods for QRS complex detection. Wavelet transform (WT) methods have been mostly used in ECG feature extraction along with other enhancements (Übeyli 2008; Kabir & Shahnaz 2012; Madeiro et al. 2012; Rezk et al. 2011; Pan et al. 2010; Martínez et al. 2004). Other techniques include geometric analysis (Zhou et al. 2009), the difference operation method (Yeh & Wang 2008), spectral analysis (Ubeyli 2008), Cumulative Sums of Squares (Alkhaldi et al. 2013) etc. However, the majority of these highly accurate QRS detection algorithms have used complex methods requiring complex time-frequency domain conversion (e.g. Fourier/Wavelet transform, complex filtering). Such methods require high computational power and memory resources, which show great limitations for real-time deployment. In order to address such issues, Elgendi et al. (2014) presented a thorough revision on QRS detection methodologies for portable, wearable, battery operated and wireless ECG devices. Multiple techniques of QRS detection from raw ECG signals were presented in their work including thresholding, syntactic methods (Lee et al. 1996; Afonso et al. 1997; Hii & Chung 2011), hidden Markov models (Borjesson et al. 1982), neural networks (Xue et al. 1992; Vijaya et al. 1998; Hu et al. 1993; Strintzis et al. 1992), matched filters (Dobbs et al. 1984; Ebenezer & Krishnamurthy 1993), singularity techniques (Virgilio et al. 1995; Rao 2015; Kadambe et al. 1999) and zero crossing (Hennig et al. 2003). Table 1 shows a summary of some of these proposed methods and their limitations. According to Elgendi et al. (2014), thresholding methods were most computationally efficient for the portable battery operated devices when conducting QRS detection. However, empirical results also indicated that the initial parameter setting for such methods proved very important in determining their performances.

Table 1 : Methods and their limitations

Method	Limitations
Thresholding	Highly dependent on initialization. Fixed threshold setting might result in performance limitation.
Neural Networks	Training phase is highly computationally intensive whereas deploying a trained neural network with a large number of connections and weights requires a considerable amount of memory.
Hidden Markov Models	Computational complexity is very high and requires a large number (15-50) of parameter evaluations.
Matched Filters	Computationally expensive due to template comparison along ECG sample by sample.

Syntactic Methods	Grammar pattern formulation and various parameter calculation are highly costly in terms of computation.
Zero-Crossing	Max/min search for temporal localization of R wave is time consuming and computationally inefficient.

The popular methods used in ECG feature extraction is discussed below.

3.2.2.1 Thresholding

Thresholding is the simplest method used in R-peak detection and feature extraction. A threshold is a value (static or dynamic) used as a determiner or decision maker. For example, consider a system which detects amplitude fluctuation of the ECG signal. If it is more than a certain value (threshold value), the system considers it as a spike. Thresholding is generally used with other methods, otherwise, it is very simple and prone to errors. It is the easiest method for implementation and takes very low computing resources as the decision (like peak detection) is made based on one or multiple comparisons of values.

3.2.2.2 Filtering

In signal processing, filtering is the process of removing unwanted signal components or features to obtain the intended part or feature of the signal. Filtering is easier to implement and works well in noise reduction. However, for feature extraction from ECG signal it is not similarly effective.

There are multiple realizations of linear filters. The following are the commonly used ones:

- **Low-pass filter:** where only low frequencies are passed, high frequencies are attenuated.
- **High-pass filter:** where only high frequencies are passed, low frequencies are attenuated.
- **Band-pass filter:** where only frequencies in a frequency band are passed.
- **Band-stop filter:** where only frequencies in a frequency band are attenuated.
- **All-pass filter:** where all frequencies are passed, but the phase of the output is changed.
- **Notch filter:** which rejects just one specific frequency.
- **Comb filter:** which has multiple regularly spaced narrow passbands resulting the appearance of a comb.

The combination of multiple filters can be used to extract important parts (features) of the signal for further processing.

3.2.2.3 Fourier Transform

Fourier transform takes a time domain signal (containing a function of time) and produces a frequency domain signal (containing frequencies) (Pinsky 2009). Fourier transform decomposes the signal into sines and cosines. The following equation shows conversion mathematically:

$$X(f)=\int_{-\infty}^{\infty} x(t)e^{-i2\pi ft}dt$$

which converts a function of time *x (t)* into a function of frequency *X (f).*

The opposite operation is called inverse Fourier transform where a function of frequency is taken and result is a time domain function. The process can be mathematically formulated as follows:

$$X(t)=\int_{-\infty}^{\infty} X(f)e^{i2\pi ft}dt$$

3.2.2.4 Wavelet Transform

Wavelet transform is very similar to Fourier transform (or closer to the windowed Fourier transform) (Pinsky 2009), however, it uses a completely different merit function which is localized in both real and frequency (Fourier) domain. The wavelet transform can be expressed mathematically using the following formula:

$$F(a,b)=\int_{-\infty}^{\infty} f(x)\,\psi^{*}_{(a,b)}(x)\,dx$$

Where ψ is some function and * is the complex conjugate symbol. As the function ψ can be chosen arbitrarily by obeying certain rules, the wavelet transform is an infinite set of various transforms.

Wavelet transforms can be categorised in two types, i.e. discrete wavelet transform (DWT) and continuous wavelet transform (CWT). A wavelet is a mathematical function that can divide a given function or continuous-time signal into diverse scale components. Generally, one can allocate a frequency range to each scale component.

In the discrete wavelet transform, orthogonal wavelets are used. The output of the discrete wavelet transform is a data vector which has the same length as the input. Typically, even in the resulting vector many data are nearly zero. This means that it decomposes into a set of wavelets (functions) which are orthogonal to its translations and scaling. Hence, we decompose such a signal to a same or lower number of the wavelet coefficient spectrum than the number

of signal data points. This type of a wavelet spectrum is very good for signal processing and compression as no redundant information is added in the output.

In the continuous wavelet transform non-orthogonal wavelets are used. The continuous wavelet transform returns one dimension array larger than the input data. For a one-dimensional data vector, an image of the time-frequency plane is obtained. The signal frequency evolution during the duration of the signal and comparison of the spectrum with other signals spectra can be easily observed. As the non-orthogonal set of wavelets is used in continuous wavelet transform, data are highly correlated, thus vast redundancy is present here.

3.2.3 ECG classification

Researchers have shown that the neural networks with different configurations were widely used for the classification of ECG signals (Ceylan et al. 2009; Özbay & Tezel 2010; Moavenian & Khorrami 2010; Babu 2013; Kumar Jaiswal & Paul 2014; Gothwal et al. 2011; Jen & Hwang 2008; Tayel & El-Bouridy 2008). There are also many other classification techniques used, most notably; support vector machine (Zhang et al. 2014), multi-layer perceptron (Magjarevic et al. 2009; Mai et al. 2011) and template matching (Magjarevic et al. 2009).

3.2.3.1 Instanced-based Algorithms

In instance-based learning a decision problem with instances or examples of training data is considered as important or required for building the model. These methods usually build up a database of example data and associate new data to the database by means of a similarity measure to find the best match and make a prediction. Instance-based methods are called memory-based learning or winner-take-all methods for this reason. The intention is to find the representations of the stored instances and conduct similarity measures between instances (Aha et al. 1991) .

The most popular instance-based algorithms are:

- k-Nearest Neighbour (kNN)
- Learning Vector Quantization (LVQ)
- Self-Organizing Map (SOM)
- Locally Weighted Learning (LWL)

3.2.3.2 Bayesian Algorithms

Bayesian methods explicitly apply Bayes' Theorem for classification and regression (Murty & Devi 2011).

The most popular Bayesian algorithms are:

- Naive Bayes
- Multinomial Naive Bayes
- Gaussian Naive Bayes
- Bayesian Network (BN)
- Bayesian Belief Network (BBN)
- Averaged One-Dependence Estimators (AODE)

3.2.3.3 Decision Tree Algorithms

In decision tree a model of decisions is made based on actual values of attributes in the data. For a given record decisions fork in tree structures until a prediction is made. Decision trees can be trained on data for classification and regression problems. Decision trees are usually fast and accurate and widely used in the machine learning domain (Safavian & Landgrebe 1991).

The most popular decision tree algorithms are:

- Iterative Dichotomiser 3 (ID3)
- Classification and Regression Tree (CART)
- Chi-squared Automatic Interaction Detection (CHAID)
- C4.5 and C5.0
- Decision Stump
- Conditional Decision Trees
- M5

3.2.3.4 Neural Network Algorithms

Artificial Neural Networks (ANN) models are proposed inspired by the functions or structures of biological neural networks. This type of pattern recognition/matching is commonly used for regression and classification problems, however can be applied on many other domains (Zhang 2000).

The most widely used artificial neural network algorithms are:

- Back-Propagation
- Radial Basis Function Network (RBFN)
- Hopfield Network

Neural networks are also used in deep learning; however, such discussion is out of scope for this thesis.

3.2.3.5 Support Vector Machines

Support vector machines (SVM) (which are also known as support vector networks) are supervised learning models with appropriate learning algorithms which are usually used for classification and regression problems (Schölkopf 1998). Given a set of training examples with labels (that marks which class the example belongs to), a support vector machine training algorithm builds a model that assigns new examples to certain categories making it a non-probabilistic binary and linear classifier. The outputs of the SVM model can be represented as points in space, mapped in such a way that the examples of the various categories are divided by a clear gap that is as wide as possible. New data examples are then mapped into the same space and the prediction is determined based on its position in the space (depending on which side of the gap the new example falls in). In addition to performing linear classification, SVMs can efficiently perform a non-linear classification using the kernel trick by implicitly mapping their inputs into high-dimensional feature spaces.

3.2.3.6 Ensemble Algorithms

Ensemble methods are models made of multiple weaker learners (base models) which are independently trained (Rokach 2010). Then, the predictions of weak learners are combined in certain ways to make the final prediction. To build an ensemble classifier most of the effort should be concentrated on (1) making diverse base learners and (2) also making a better mechanism for combining decisions of the weak learners for the improvement of final accuracy. Some of the popular ensemble techniques are listed below.

- Boosting
- Gradient Boosting Machines (GBM)
- Bootstrapped Aggregation (Bagging)
- AdaBoost
- Random Forest
- Stacked Generalization (Blending)
- Gradient Boosted Regression Trees (GBRT)

3.2.4 Novel class detection

Novel class detection is used in many domains and can be applied in ECG analysis for unknown beat detection. The novel class detection approaches are studied in detail by Markou and Singh

(2003a). They divided the approaches in two parts. The first part explains all the statistical approaches (Markou & Singh 2003a). In that case modelling is conducted based on statistical properties and that is used to find out whether a test sample belongs to the same distribution or not. As most of the work is done by modelling data distribution and probability estimation of test data, training data quality becomes essential for robustness of the classifier. The second part reviews most of the neural network based techniques (Markou & Singh 2003b). Hou et al. (2010) proposed 'RedTrees' which is developed based on a decision tree for multi-relational data streams classification where every node is considered as one atom and the split manners were selected by information entropy. Masud et al. (2011) proposed a novel class detector, i.e. 'ECSMiner' (Masud et al. 2011), using traditional methods. Their method collects multiple test instances to predict whether they belong to novel class or not. It is a supervised method for making a boundary of existing classes in training dataset. The trained classifier was able to identify outliers and declare novel classes. Based on this, MCM (Multi Class Miner) was proposed which works with feature-evolving data streams (Masud et al. 2013). A feature set homogenization technique was introduced to address feature-evolution and to make novel class detection more precise by creating adaptive boundary. In another work (Miao et al. 2013), the performance of 'ECSMiner' is improved by changing base learners from normal decision trees to Very Fast Decision Trees (VFDTs) integrated with clustering algorithms. Elwell and Polikar (2011) proposed 'Learn^{++} NSE' which is incremental ensemble classifier. It trains a new classifier for each new batch of data and they are combined based on a dynamic majority voting strategy. Classifier weights are adjusted by its accuracy on current and past environments. A cluster based approach was proposed by Spinosa et al. (2008). It is a bi-class classifier differentiating between normal and novel class only. Katakis et al. (2006) proposed an incremental feature ranking method along with a learning algorithm. It is another statistical method which uses either k-NN or Naive Bayes. Another incremental method 'FAE' is proposed by Wenerstrom et al. (2006). In their model, an ensemble is used instead of single incremental learner. In this case, a vector of top contributing features is kept and the experts are updated frequently discarding obsolete models. The final decision is taken by majority voting. Another interesting ensemble classifier with novel class detection is proposed by Masud et al. (2011). Their method tries to create boundaries for each model while training. Any test instance falling outside is considered an outlier which is stored for cohesion check. When multiple instances have enough cohesion and are different from the trained classes, they are considered as a novel class instances. The problem with this approach is that a real-time prediction is not feasible as it needs to store a buffer of outliers rather than predicting on each

arriving instance. The second problem is if there are multiple novel classes present, the method will fail to find enough cohesion among the instances belonging to them. Farid and Rahman (2012) proposed a classifier using a customized decision tree algorithm with the help of the clustering technique to detect novel classes from data streams and recently another improved version is proposed by the authors which involves creating ensembles of decision trees (Farid et al. 2013). In another recent work of ZareMoodi et al. (2015), novel class in data stream is handled by creating ensembles per each class basis. The oldest cluster based novelty detection method 'OLINDDA' (Spinosa et al. 2007) was proposed in 2007. It is completely unsupervised and the threshold based single class novelty detection was conducted using the K-Means clustering algorithm. Later on an improved version 'MINAS' (Faria et al. 2013) was proposed. It has similar training and detection procedures but a multi-class approach is taken.

3.2.5 Optimization

There are many optimization algorithms available in the literature, however, this research is mainly focused on Swarm intelligence (SI) based optimization as these methods have gained much popularity recently (Kennedy et al. 2001). Popular SI algorithms include Particle Swarm Optimization (PSO), Genetic Algorithm (GA) (Holland 1975), Cuckoo Search Optimization (CSO) (Yang & Deb 2009), Artificial Bee and Ant Colony Optimization. In each of the swarm based optimization algorithms, the movement of the agents are controlled by the behaviours of a natural phenomenon. In general, increasing the number of search agents results in higher diversity and reaching promising optimal solutions in fewer iterations. Meanwhile, increasing the dimension of the problem domain makes the optimization problem more challenging to reach global optima.

3.2.5.1 Firefly Algorithm

The firefly algorithm (FA) introduced by Xin-She Yang (Yang 2009) is inspired by movement of fireflies based on its bioluminescence. It employs the following three rules to guide the search process. (1) Fireflies are unisex and attracted to each other. (2) The attraction is proportional to brightness and inversely proportional to the distance between two fireflies. As a result, the less bright fireflies will move towards the brighter fireflies. The brightest firefly will move randomly. (3) The brightness of each firefly represents the solution quality. Initially, the algorithm generates a population of fireflies randomly. Subsequently, the fitness value of each firefly is calculated. Fireflies are ranked according to their fitness. Then the following

equation is applied to move a firefly X_i with a lower light intensity towards a brighter one, X_j, in the neighbourhood.

$$X_i^{t+1} = X_i^t + \beta_0 * e^{-\gamma r_{ij}^2}\left(X_j^t - X_i^t\right) + \alpha_t \varepsilon_t \qquad (3.1)$$

where β_0 denotes attractiveness when distance $r = 0$. r_{ij} is the distance between firefly i and firefly j. γ is the fixed light absorption coefficient. α_t denotes the randomization parameter that controls the step size of the randomized move and ε_t is a random walk behaviour defined by Gaussian or other distributions. The above search steps are repeated until the algorithm reaches the maximum number of iterations.

3.2.5.2 Other FA Variants

Recently, multiple modifications on FA are proposed. For example, a FA variant was proposed by Husselmann et al. (2012) which utilized penalization so that it can be used efficiently on graphic processing units (which contain array of parallel processors). Their algorithm achieved higher accuracy than that obtained by the original FA for multi-modal optimization functions. Another FA variant, known as parallel FA, was developed by Subotic et al. (2012) for unconstrained optimization problems. The test results showed the efficiency of the proposed algorithm when evaluated using standard benchmarking functions. Tilahun et al. (2012) proposed Elitist Firefly which modified the random movement of the brightest firefly by decreasing its illumination when the current best position did not improve after a certain number of generations. It improved the brightest firefly by generating m-uniform random vectors and moving it towards the global optima. Experiments with seven benchmarking functions indicated that it outperformed the original FA significantly. Gandomi et al. (2013) integrated chaotic map with FA to increase its global search mobility. Another attempt of combining chaotic map and the FA was conducted by Coelho et al. (2011). Their proposed algorithm had better capabilities to escape from local optima to achieve global convergence. Gaussian distribution was used by Farahani et al. (2011) in their research to modify the randomization parameter α of the FA adaptively to balance between local exploitation and global exploration. Lévy flight was also used in other FA modification proposed by Yang (Yang 2010; Yang 2012) as the random distribution to increase global exploration.

3.3 Conclusion

Wearable technology and remote monitoring seem to be the future of the healthcare as it significantly reduces the barrier between patients and doctors. With the growing population, the workload on healthcare is increasing. For those living in remote locations or for someone

who needs continuous monitoring, wearable, wireless mobile monitoring systems could provide the best possible solution in absence of physicians. The mobile ECG platforms are growing rapidly, resulting in multiple new commercial systems in recent years. However, many software and hardware improvements can be made on existing systems.

This research will mainly focus on software improvements by developing novel methods for ECG processing that can be easily used on mobile systems or can be implemented in cloud servers for further processing. As the ECG processing expands into multiple sub-domains of computer science (such as digital signal processing, optimization, and machine learning), there could be a variety of improvements in each domain, resulting in better outcomes of ECG analysis in terms of computational cost and abnormality detection accuracy. The proposed methods of each key stage are discussed in each of the subsequent chapters in this thesis.

Chapter 4 ECG Pre-processing

In this thesis, multiple algorithms are proposed to improve the overall accuracy and performance of automated ECG processing. The contributions are distributed over multiple stages as illustrated in Figure 6. In the ECG pre-processing stage, a baseline correction algorithm is proposed. Subsequently, an offline noise reduction method is also developed at this stage. Both mentioned algorithms are discussed later in this chapter. Three algorithms (3, 4, and 5 in Figure 6) are proposed for ECG feature extraction which is discussed in detail in Chapter 5. A feature enhancement strategy is developed and described in Chapter 6. For the classification stage, a hybrid ensemble classifier is proposed along with a novel class detector

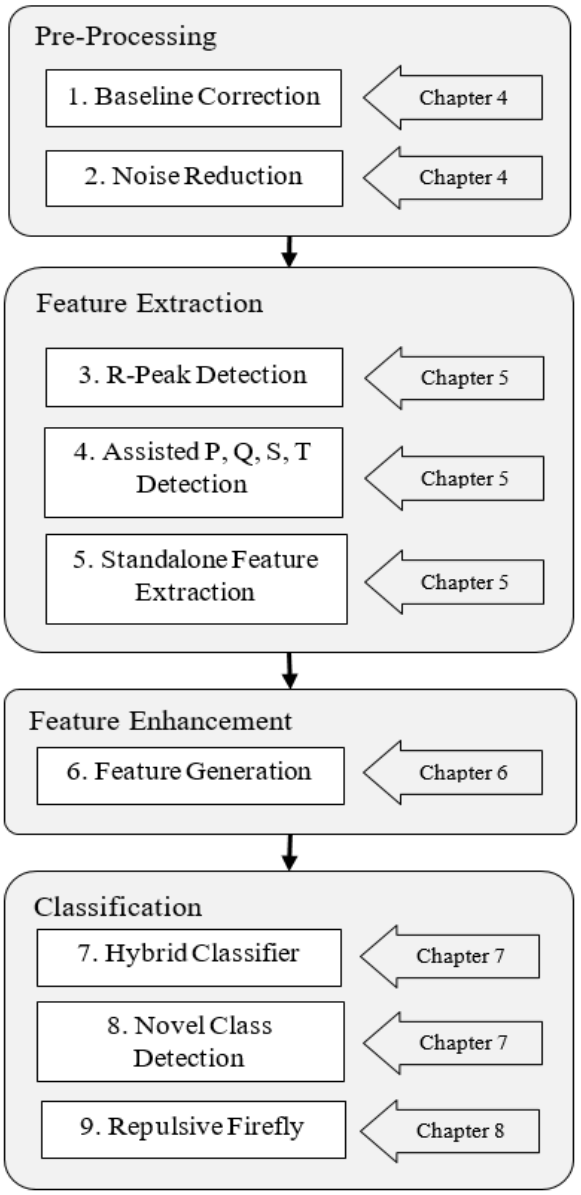


Figure 6 : The proposed ECG processing system.

(7, 8 in Figure 6). They are discussed in detail in Chapter 7. Parameters of both of the classifiers are set by an evolutionary optimization algorithm (9 in Figure 6) which is presented in Chapter 8. The rest of this chapter discusses the contributions in ECG pre-processing stage. Raw ECG signals taken from ECG sensors might include several types of noise which is discussed earlier in the pre-processing section of Chapter 2. In this chapter, a noise reduction technique is proposed which can be applied to reduce noise artefacts in ECG signals acquired from both the mobile and traditional ECG sensors. This algorithm can be used for offline processing on a pre-recorded ECG signal. The proposed algorithm uses stationary wavelet transform (Nason & Silverman 1995) and dynamic thresholding to deal with the removal of various types of noise. The evaluation results of noisy ECG signals also demonstrate the effectiveness of the proposed algorithm. The rest of the chapter is divided into three sections: Method, Experimental Results, and Conclusion. The method section describes the details of the algorithm. The experiment section elaborates on the experiments conducted and presents the results. Finally, the findings are summarized in the conclusion section

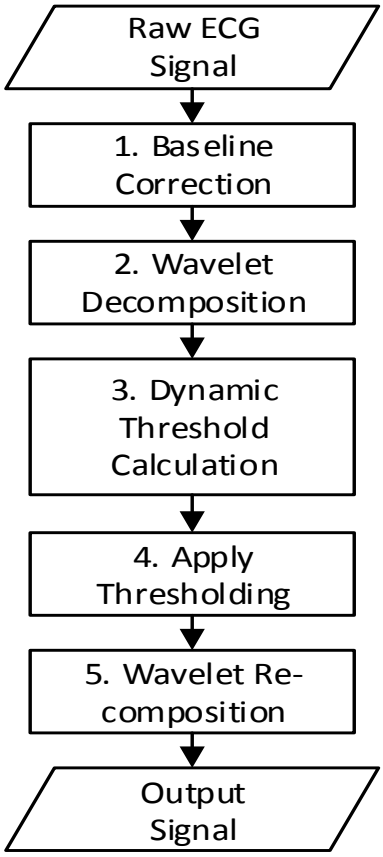


Figure 7 : Flowchart of the proposed noise reduction system.

4.1 Method

In this section, a noise reduction technique is proposed which incorporates five main steps as illustrated in Figure 7. In the first step, the baseline is corrected using moving window average. Then, the signal is decomposed using discrete stationary wavelet transform. From the decomposed signal, the thresholds are then calculated dynamically in the third step. Next, the calculated thresholds are applied to the decomposed signal. Finally, the output signal is recomposed using inverse discrete stationary wavelet transform.

4.1.1 Baseline Correction

In this stage, any baseline related deviations of the ECG signal are corrected. This is a necessary step as the subsequent dynamic threshold generation will be affected if the signal refers to an incorrect baseline. A simple baseline removal technique is proposed which subtracts the averaging window (around the current sample) of half a second from the current sample to remove baseline deviations. The operation is defined as follows:

$$y_b(x) = \begin{cases} y(x) - \frac{1}{w}\sum_{i=1}^{w} y\left(x + i - \frac{w}{2}\right) & If\ \frac{w}{2} \le x \le size(y) - \frac{w}{2} \\ y(x) & Otherwise \end{cases} \quad (4.1)$$

where $w = round(fs \times 0.5) + round(fs \times 0.5)\%2$ (4.2)

where *y(x)* denotes the raw ECG signal (with a baseline problem) and y_b *(x)* denotes the corrected signal. *fs* is sampling frequency and *size*(*y*) returns number of samples in *y(x)*. The window size *w* is an even number which is close to the number of samples in half a second. An example is illustrated using this technique in Figure 8, where the raw signal is shown on top of the diagram and the corrected signal is plotted at the bottom.

4.1.2 Wavelet Decomposition

A stationary wavelet transform (SWT) is used in this stage. A detailed discussion of SWT can be found in the original literature and work from Nason & Silverman (1995). The decomposition process depends on the type of wavelet family used and on the decomposition level *N*. The output of the decomposed signal contains both the detailed coefficients for each level (till *N*) and the detailed approximation coefficients of level *N*. A level 3 decomposition result is shown in Figure 9. The noisy (but baseline corrected) signal (i.e. signal a) is illustrated in the top row of Figure 9. The next three signals (b, c and d) in Figure 9 are resulting decomposed signals with the last row representing the filtered signal (i.e. signal e).

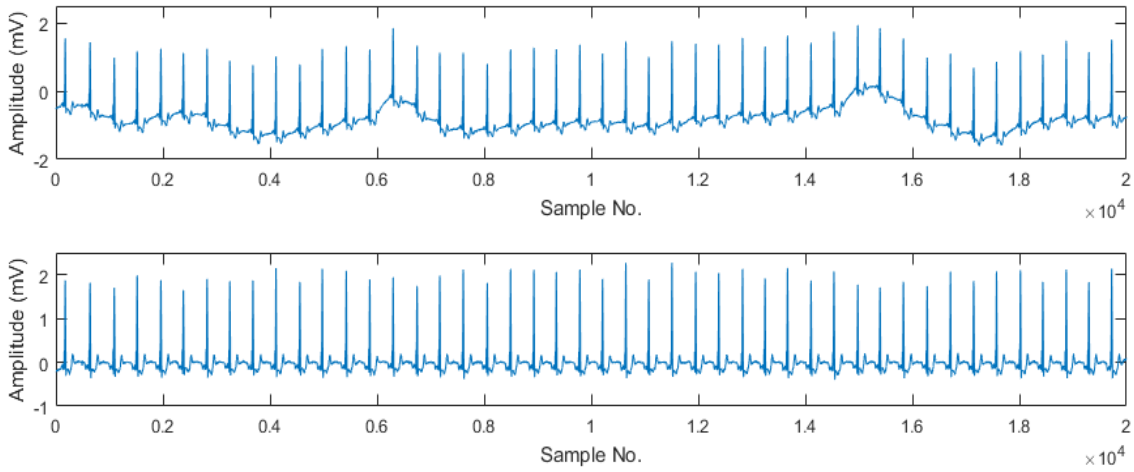


Figure 8 : An ECG signal with baseline problem and a baseline corrected signal.

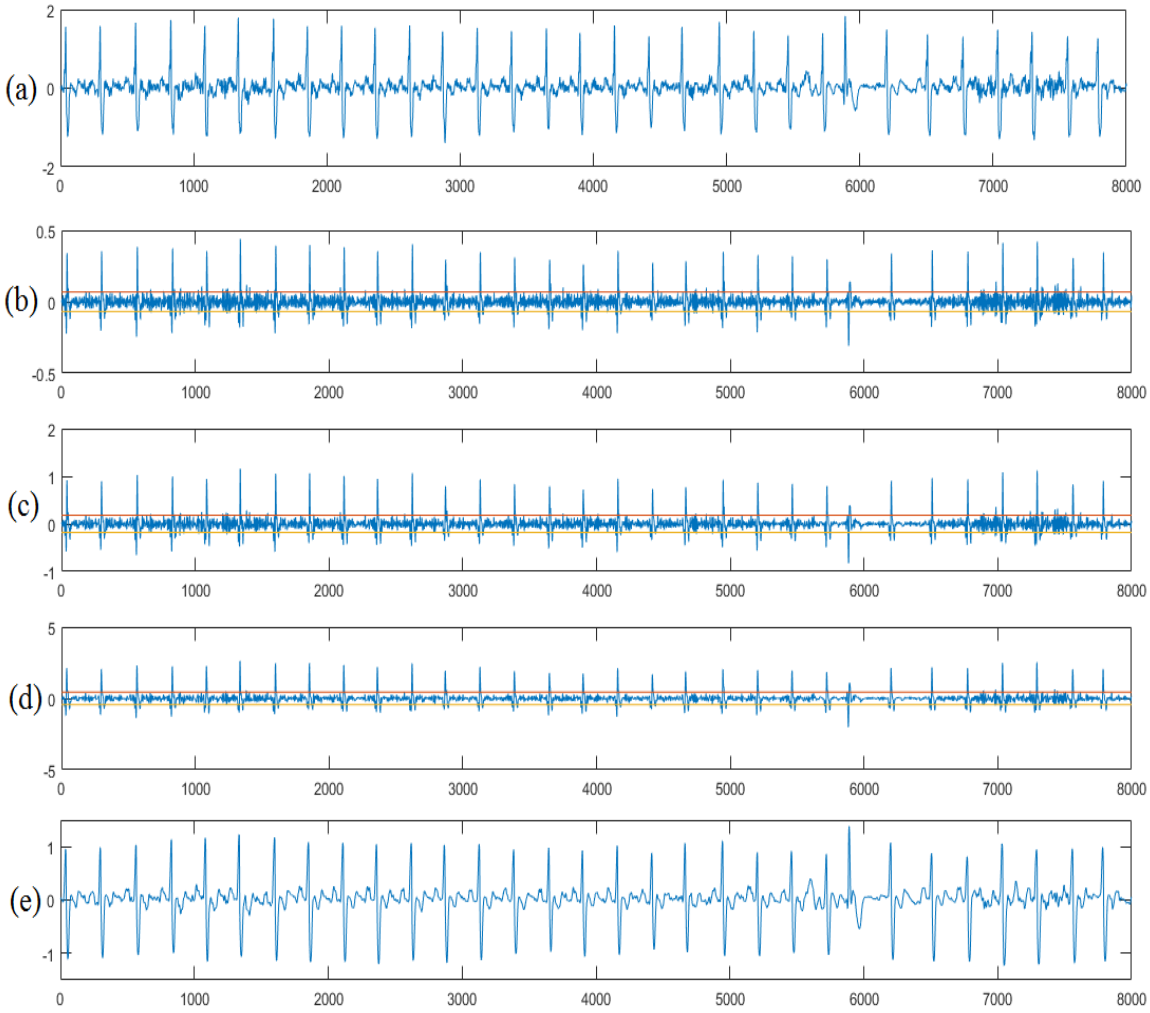


Figure 9 : (a) Noisy signal, (b)/(c)/(d) decomposed signal and (e) the filtered signal.

4.1.3 Threshold Calculation

In this step, the amplitude approximation thresholds are calculated for each decomposed signal. The threshold is calculated using a certain percentage value *tp*. This indicates only about *tp*% of sample values exist, and are above the output cut-off threshold *ta*. The algorithm for cut-off threshold calculation from given *tp* is discussed in detail in the following.

Algorithm 1: Cut-off threshold calculation

Input:

tp //Percentage of values above present threshold

y (x) //The decomposed ECG signal (after baseline correction)

Output:

ta //Cut-off threshold for the decomposed ECG signal

Begin

{

//Find the average of the mean of the absolute signal and the peak of the signal as the starting threshold

$t = [(\sum_{i=1}^{L} absolute\ (y\ (i))\ )/L + \max\ (y)\ ]/2$ //L is the number of elements in the signal

dr = 0.001; //Decrement amount

s = true;

While s = true **do**

t = *t* – *dr* ;

n =$\sum_{i=1}^{L} gt\ (absolute\ (y\ (i))\ , t)$; // *gt* (*x*,*t*) returns 1 if *x* > t; otherwise it returns 0.

If *n*/*L* > *tp*

s = false;

End

End

ta = *t*;

Output *ta*

}

End

In above algorithm cut-off threshold adapts to a value for the given signal so that only *tp*% of the samples are discarded (cut off). In Figure 9, an example is provided where threshold values are calculated using *tp* = 10%. The resulting *ta* is marked as positive and negative straight lines for each decomposed signal (b, c and d).

4.1.4 Threshold Application

Next, the calculated threshold is applied to all of the decomposed signals. The following formula is used for the application of the threshold on the corresponding decomposed signal.

$$y_a(x) = \text{sign}\,(y(x)) \times \text{pos}\,(absolute(y(x)) - T) \quad (4.3)$$

where sign (*x*) returns 1 if *x* is positive and returns -1 otherwise. *pos*(*x*) returns *x* if the value of *x*>0; otherwise it returns 0. y_a represents the output signal after applying the threshold on the input signal *y*. In Figure 9, (b, c and d), threshold values are marked as two straight lines along the x-axis (for each signal) with both having the same positive and negative amplitudes. All the values within the threshold regions will be reduced to 0.

4.1.5 Wavelet Re-composition

The filtered decomposed signals are used in this stage to reconstruct the final filtered ECG signal. This step involves inverse discrete stationary wavelet transform. Detailed inner workings of this process can be found in (Nason & Silverman 1995).

4.2 Experimental Results

In this section, multiple experiments using different families of wavelets are conducted. Various levels of wavelet decomposition are performed for different generated noisy signals with each having increasing signal-to-noise ratio (SNR) than the previous signal. All of our experiments are conducted using MATLAB software.

4.2.1 Experimental Data

A single real ECG signal, 'signal 100' from MIT-BIH Arrhythmia (Moody & Mark 2001) database is used as the base signal. Different types of noise were added to the base waveform, which result in varied SNR for each of the generated test signals. Added noise includes both 60Hz sinusoidal signal and Gaussian white noise. A set of 20 noisy signals are generated by integrating time varying and amplitude varying noise signals with the original ECG signal (i.e. signal 100). Each of the generated noisy signals has increasing noise power, resulting in SNR starting from -0.0182 dB up to -4.2091 dB. The base waveform contains a half-hour-long real ECG recording sampled at 360Hz. Each of the generated noisy signals is of the same length containing different power and variation of noise. A summary of increasing SNR for all twenty generated noisy test signals can be found in Table 2.

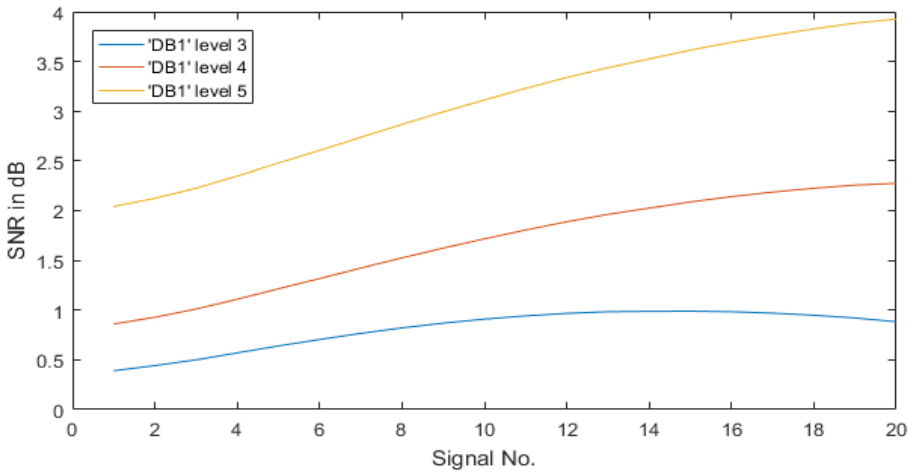


Figure 10 : Comparison of the SNR of the output signals using multiple levels of 'DB1' wavelet transform

Table 2 : Summary of SNR in dB for all the test signals

Signal No.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
SNR	-0.01	-0.06	-0.14	-0.24	-0.38	-0.53	-0.71	-0.91	-1.12	-1.35	-1.58
Signal No.	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
SNR	-1.35	-1.58	-1.83	-2.08	-2.34	-2.60	-2.86	-3.12	-3.39	-3.65	-3.91

4.2.2 Evaluation Results

The proposed algorithms are tested using 'Daubechies' family of wavelets (DB2, DB3 and DB4) and level 3, 4 and 5 decompositions. Two sets of experiments were conducted. In the first experiment, *tp* is considered as 10%. Test results of the first set are shown in Table 3. To enable visual comparison, two diagrams are plotted which are shown in Figure 10 and Figure 11, respectively. In Figure 10, SNR outputs using the same family of wavelet ('DB1') with varying decomposition levels for each of the test signals are being compared. It seems that the increasing decomposition level greatly improves the final SNR output. In Figure 11, the outputs, using the same level of decomposition but different families of wavelet for each signal, are compared. Level 1 decomposition is selected for this case as higher level decomposition experiments are done in later stagesThe test result shows that the outputs from different families of wavelet (for the same level of decomposition) vary for the signals with low noise. However, they tend to produce comparable results when a high amount of noise is present.

Table 3 : Comparison of the final SNR of the output signals using the proposed system with *tp* =10%

Wavelet Family >		DB1			DB2			DB3			DB4			DB5		
Signal No.	Initial SNR	Decomposition Level			Decomposition Level			Decomposition Level			Decomposition Level			Decomposition Level		
		3	4	5	3	4	5	3	4	5	3	4	5	3	4	5
1	-0.01	0.39	0.86	2.04	0.46	1.39	2.60	0.54	1.96	3.28	0.76	2.37	3.96	0.79	2.49	4.16
2	-0.06	0.44	0.93	2.12	0.50	1.44	2.66	0.56	1.99	3.34	0.77	2.40	4.00	0.81	2.52	4.21
3	-0.14	0.50	1.01	2.23	0.53	1.48	2.72	0.59	2.04	3.40	0.79	2.42	4.04	0.83	2.56	4.27
4	-0.24	0.57	1.11	2.35	0.57	1.53	2.80	0.61	2.07	3.44	0.79	2.44	4.07	0.84	2.59	4.33
5	-0.38	0.64	1.21	2.48	0.61	1.59	2.88	0.63	2.10	3.50	0.79	2.46	4.10	0.85	2.62	4.37
6	-0.53	0.70	1.32	2.61	0.65	1.66	2.98	0.65	2.14	3.57	0.79	2.46	4.12	0.85	2.63	4.41
7	-0.71	0.77	1.42	2.74	0.69	1.72	3.07	0.67	2.18	3.64	0.79	2.48	4.15	0.85	2.65	4.45
8	-0.91	0.82	1.53	2.87	0.72	1.78	3.16	0.69	2.22	3.71	0.79	2.49	4.19	0.84	2.66	4.48
9	-1.12	0.87	1.62	2.99	0.75	1.83	3.24	0.71	2.26	3.78	0.79	2.51	4.23	0.83	2.67	4.52
10	-1.35	0.91	1.72	3.11	0.76	1.88	3.31	0.72	2.28	3.83	0.78	2.52	4.28	0.83	2.69	4.56
11	-1.58	0.94	1.81	3.23	0.77	1.92	3.39	0.72	2.31	3.90	0.77	2.54	4.32	0.81	2.70	4.60
12	-1.83	0.97	1.89	3.34	0.78	1.96	3.45	0.72	2.33	3.95	0.76	2.55	4.36	0.80	2.70	4.63
13	-2.08	0.98	1.96	3.44	0.77	1.99	3.51	0.71	2.34	3.98	0.74	2.54	4.38	0.77	2.70	4.66
14	-2.34	0.99	2.03	3.53	0.75	2.02	3.57	0.69	2.34	4.02	0.71	2.54	4.41	0.74	2.69	4.68
15	-2.60	0.99	2.09	3.61	0.73	2.04	3.62	0.67	2.34	4.05	0.68	2.53	4.42	0.71	2.68	4.69
16	-2.86	0.98	2.14	3.69	0.70	2.06	3.67	0.64	2.33	4.07	0.64	2.51	4.43	0.67	2.66	4.69
17	-3.12	0.97	2.19	3.76	0.67	2.07	3.71	0.60	2.32	4.09	0.59	2.49	4.42	0.62	2.63	4.68
18	-3.39	0.95	2.23	3.83	0.63	2.08	3.74	0.55	2.30	4.10	0.54	2.46	4.42	0.56	2.59	4.67
19	-3.65	0.92	2.26	3.89	0.58	2.07	3.77	0.50	2.28	4.11	0.49	2.43	4.41	0.50	2.56	4.65
20	-3.91	0.88	2.28	3.93	0.53	2.07	3.80	0.45	2.26	4.11	0.42	2.39	4.39	0.44	2.51	4.62

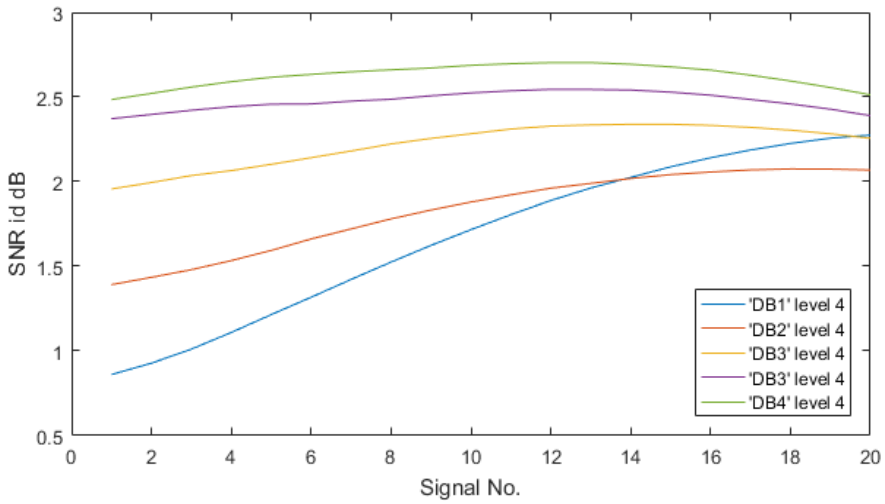


Figure 11 : Comparison of the SNR of the output signals using multiple families and level 4 wavelet decomposition

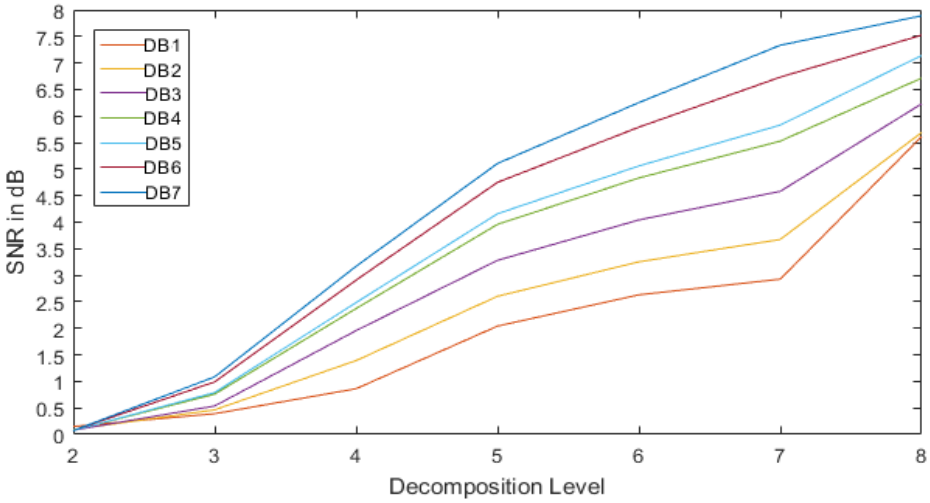


Figure 12 : Comparison of the SNR of the output signal using multiple families and levels of wavelet transform using the test signal 1.

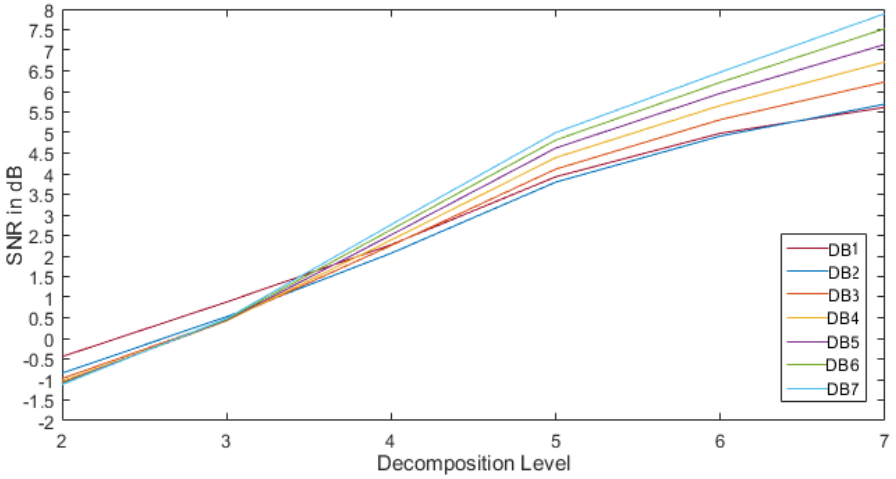


Figure 13 : Comparison of the SNR of the output signal using multiple families and levels of wavelet transform using test signal 20.

In the second set of experiments, the noisiest signal (test signal no. 20) and the least noisy signal (test signal no. 1) are taken for experimentation. Wavelets decomposition on multiple levels (level 2 to level 7) are applied, each using DB1 to DB7 families of wavelets on both signals. A detailed comparison of the test results is shown in Table 4. Visual comparisons are also provided in Figure 12 and Figure 13, respectively, for test signal 1 and test signal 20. For the input signal with low noise (Signal 1), the performance of our algorithm increases corresponding to a higher family and level of decomposition (as shown in Figure 12). This is

also the case for the processing of noisier signals, however, the improvement of SNR is comparatively less compared to the previous situation (see Figure 13).

Table 4 : Comparison of the final SNR (in dB) of output of our algorithm using two noisy signals individually

	Decomposition Levels (Signal 1)						Decomposition Levels (Signal 20)					
Wavelet Family	**2**	**3**	**4**	**5**	**6**	**7**	**2**	**3**	**4**	**5**	**6**	7
DB1	0.151	0.391	0.861	2.042	2.631	2.926	-0.444	0.883	2.275	3.927	4.983	5.614
DB2	0.092	0.464	1.392	2.603	3.253	3.674	-0.848	0.527	2.069	3.797	4.916	5.691
DB3	0.080	0.536	1.957	3.281	4.042	4.582	-0.980	0.447	2.256	4.113	5.315	6.231
DB4	0.077	0.759	2.372	3.959	4.833	5.527	-1.043	0.424	2.390	4.392	5.659	6.716
DB5	0.075	0.790	2.485	4.159	5.060	5.829	-1.081	0.440	2.513	4.624	5.954	7.142
DB5	0.076	0.991	2.908	4.750	5.790	6.733	-1.105	0.460	2.641	4.819	6.220	7.525
DB6	0.074	1.085	3.168	5.106	6.254	7.334	-1.124	0.484	2.764	5.000	6.468	7.890

4.3 Conclusions

In this chapter, an effective noise reduction technique is proposed for ECG signals using baseline removal, discrete stationary wavelet transform and dynamic thresholding. Results indicate that the efficiency of the algorithm depends on the wavelet family which is used and the applied decomposition level. The proposed algorithm is tested using a half-hour-long real ECG signal from MIT-BIH Arrhythmia database along with the generated noisy signals. The proposed algorithm can also be applied to any other one-dimensional signals (e.g. EEG signals) for white noise removal. However, if it is not an ECG signal, the baseline removal method will not be necessary. For online processing of ECG signals, only baseline removal and signal smoothing could be useful which is further explored in the following chapter. This research work was published in “Decision - Emerging Trends in Neuroengineering and Neural Computation” (Pandit et al. 2017).

Chapter 5 ECG Feature Extraction

Feature extraction is a vital part of ECG analysis since in this stage, the vital information is extracted from the ECG signal, which is used for further processing. In this chapter, two algorithms for feature extraction are proposed. Both algorithms work on raw lead II ECG signals to extract features. The first algorithm finds heartbeat locations in the raw ECG signal. This information can be used directly for heart rate variability analysis or can be used to help other feature extraction algorithms for further processing. The second algorithm is fully capable of finding P, Q, R, S, and T locations from raw ECG signals, however, heartbeat detection is not as accurate as with the first algorithm. The algorithms are discussed in detail in the following.

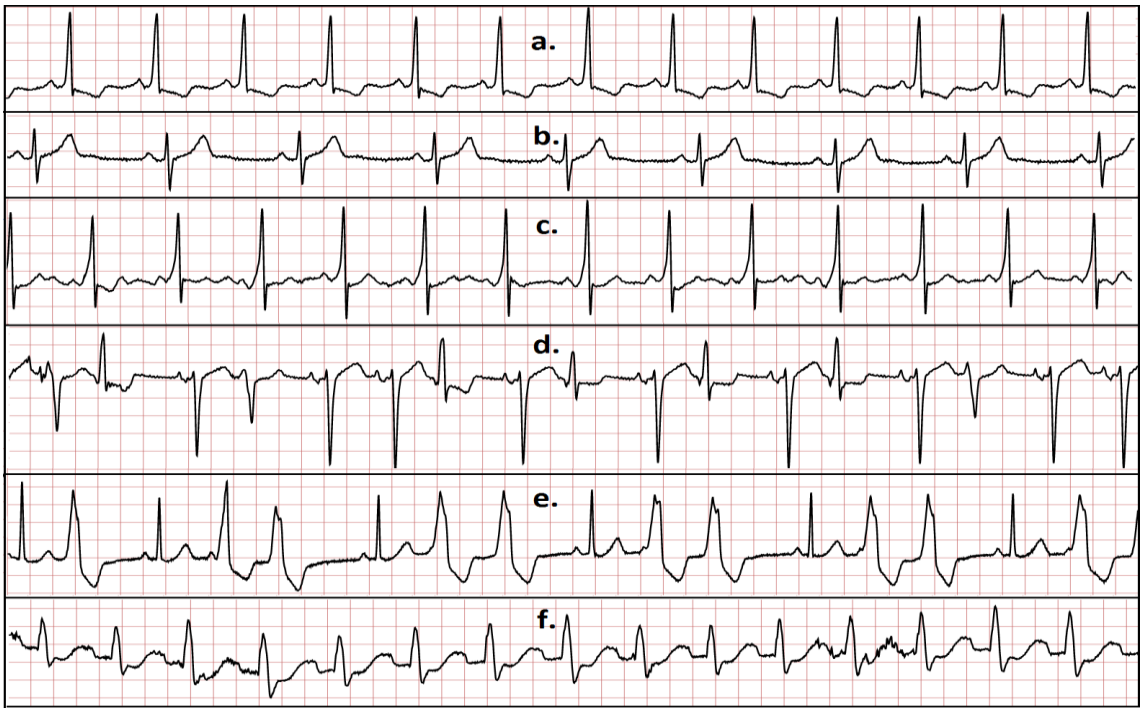


Figure 14 : Normal and abnormal real ECG signals

5.1 R-Peak detection algorithm

The detection of the R-peak in the QRS complex of an ECG signal plays a key role for the diagnosis of a patient's heart condition. To accurately identify QRS locations from raw ECG signals, many issues needed to be dealt with, for example, noise present in the signal, such as baseline wander, varying peak amplitudes and signal abnormality. In this section, a lightweight algorithm for QRS (i.e. R-peak) detection from raw ECG signals is developed which efficiently

handles the problems and locates QRS positions. The rest of this section is divided into three subsections: Challenges, Methodology and Experimental Results which are discussed below.

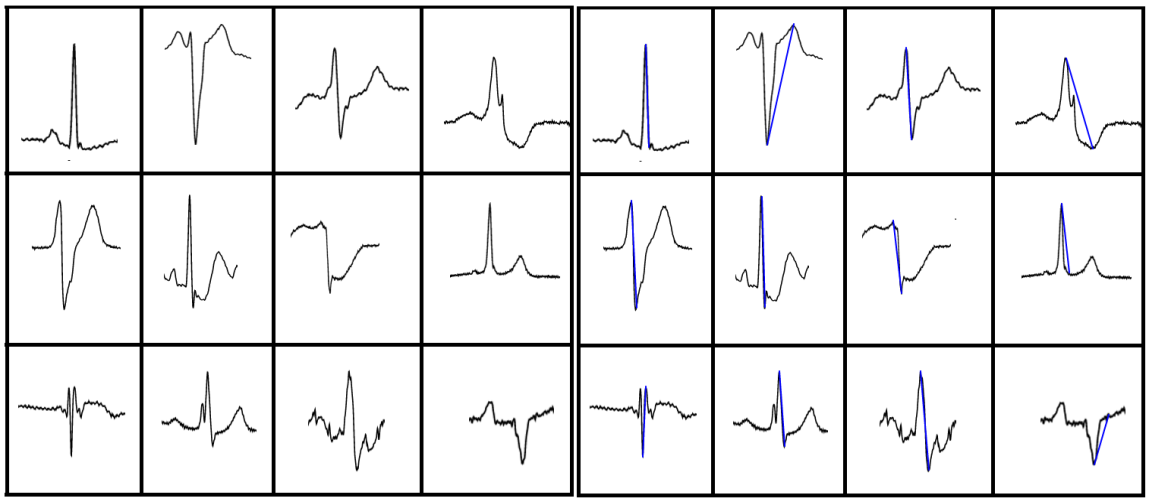


Figure 15 : (a) Series of heterogeneous QRS patterns, (b) Series of heterogeneous QRS patterns with regional highest difference of amplitudes marked (blue)

5.1.1 Challenges

QRS detection is a challenging task for mobile ECG sensors as these sensors mostly use single lead ECG signals rather than the traditional twelve lead ECG signals from a full medical ECG. Signal abnormality poses another challenge for QRS detection since the QRS patterns from abnormal ECG signals could be extremely irregular. For instance, an ideal Lead II ECG signal usually consists of P, Q, R, S and T waves. Figure 14 illustrates examples of real-world raw Lead II ECG samples from MIT-BIH Arrhythmia database (Moody & Mark 2001). As clearly shown in Figure 14, even normal patterns (i.e. a, b, and c signals in Figure 14) bear differences since the ECG signals could be influenced by each individual person's physiological conditions. Furthermore, abnormal patterns (i.e. d, e and f signals in Figure 14) can hardly be comparable to the ideal waveforms as some of the waves may be missing (e.g. the R wave is missing resulting in a QS wave in signal d). Additionally, other waves with high amplitudes (e.g. the T wave) can also be misidentified as the R wave or a QRS pattern. Figure 15 (a) also shows some detailed normal and abnormal QRS patterns taken from MIT-BIH Arrhythmia database. These QRS patterns look completely different from one another and tend to confuse state-of-the-art QRS detection algorithms. It can be observed that in most cases, the differences of amplitudes of either QR/RS waves or ST waves indicate the most dramatic change in every QRS complex. Figure 15 (b) shows the corresponding marked regions of the waves illustrated

in Figure 15 (a). These regions represent the difference between the highest and lowest points in corresponding QRS complex. This research aims to accurately identify the location of the R wave (if present) from pre-processed raw ECG signals by exploiting such features as illustrated in Figure 15 (b) to deal with signal abnormality.

Overall, this research proposes a sliding window based Max-Min Difference (MMD) approach for robust real-time QRS or R-peak detection from raw single lead (Lead II) ECG signals with efficient computational cost. Specifically, it employs dynamic thresholding along with Max-Min difference curve generation for QRS pattern detection. The proposed MMD algorithm for R-peak detection is evaluated with ECG signals extracted from multiple databases including MIT-BIH Arrhythmia (Moody & Mark 2001), European ST-T (Taddei et al. 1992), MIT-BIH ST Change (Albrecht 1983), St.-Petersburg Institute of Cardiological Technics 12-lead Arrhythmia (Anon 2016) and QT (Laguna et al. 1997) databases. To ascertain the system's efficiency in dealing with abnormality detection, the proposed QRS detection algorithm has also been combined with a feature extraction process and a neural network classifier to conduct abnormal heartbeat detection via the evaluation of the MIT-BIH Arrhythmia database.

The main contributions in this part of research work are as follows: (1) A sliding window based strategy is employed for online real-time ECG analysis. Instead of requiring the whole signal to be stored in memory, it works sufficiently with a small buffer of the signal to save computational cost. (2) A novel algorithm (i.e. MMD) is proposed for robust QRS detection. MMD calculates the difference between the local minimum and maximum of a small window and slides it along the signal to find QRS complex locations. In comparison to related research, it has better trade-off between speed and accuracy and shows impressive performances with great computational simplicity. (3) Dynamic thresholding is also proposed to deal with the fluctuating average peak amplitudes in ECG signals. It calculates thresholds by using the current peak location, a few previous QRS locations and the distances between them. (4) The proposed approach is evaluated with five ECG databases for QRS detection. A comprehensive evaluation is also conducted for the well-known MIT-BIH Arrhythmia database in comparison to state-of-the-art algorithms for R-peak detection. To further indicate the superiority of the proposed MMD algorithm, abnormal/normal heartbeat detection based on the QRS detection is also conducted by incorporating with a feature extraction algorithm and a neural network classifier. Overall, the empirical results indicate the superiority of our algorithm over other methods in terms of system performance and computational efficiency.

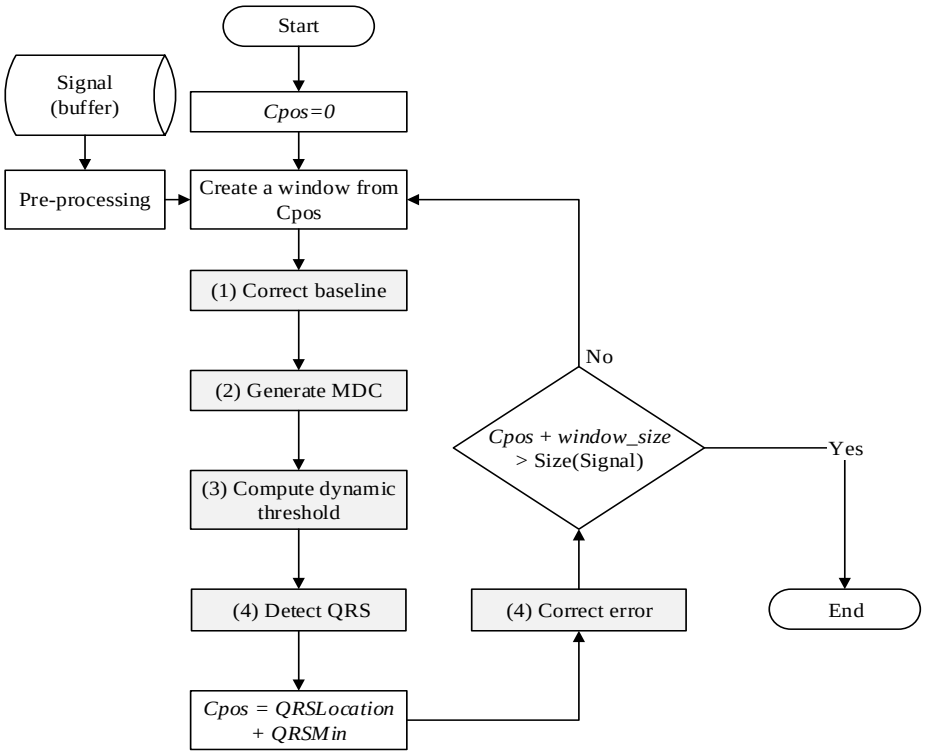


Figure 16 : Flowchart of the proposed QRS detection component

5.1.2 Methodology

In this section, the proposed approach is introduced along with some discussions on the main challenges encountered while detecting QRS complexes or R-peak locations. The main hurdles to overcome ensuring high detection accuracy usually include noise, baseline wander, varying thresholds, and signal abnormality. The corresponding proposed strategies in this research to solve the above challenges are discussed below.

- First, to deal with noise embedded in ECG signals, instead of using any band pass filtering methods, sliding window averaging is used for smoothing the signal for QRS detection and analysis. Although this process scales down all the peaks, it does not affect the accurate detection of QRS patterns. Additionally, sliding window average is not only effective, but also computationally efficient.
- Secondly, rapid changes in the baseline of the ECG signal, i.e. baseline wander, (see Figure 17) could have a negative impact on the overall detection accuracy. To deal with this problem, a simple baseline correction algorithm is used which involves window based averaging and deducting the local average of the signal from the raw signal.

- Thirdly, finding an optimal threshold for peak detection is another challenging task as the amplitude difference between R wave and P or T waves varies not only for signals collected from different subjects but also for those gathered from the same subject. Therefore, dynamic thresholding is proposed in this research to identify an efficient threshold (i.e. the minimum value which is above all P or T waves) in diverse cases.
- Another challenge is signal abnormality. E.g. sometimes R waves could go below the peak detection threshold or be missing (e.g. the d signal in Figure 14) which makes accurate detection of QRS positions significantly difficult. As an example, an abnormal heartbeat sometimes might not have an R wave present. In such cases, either a P wave or a T wave can be mistaken as an R wave. Even though an R wave is present, sometimes, P and T waves can have higher amplitude than the R wave. Finally, varying R-R interval is also a hurdle in QRS analysis. Because of unpredictable R-R intervals in abnormal signals with similar R and T wave amplitudes, it often makes peak detection algorithms fail to distinguish between them. Figure 15 (a) shows some of the QRS patterns taken from MIT-BIH Arrhythmia database which look completely different from one to another and tend to confuse many pattern recognition algorithms. It is observed that in most cases, the differences of amplitudes of either QR/RS waves or ST waves indicate the most dramatic change in every QRS complex. Figure 15 (b) shows the corresponding marked regions of the waves illustrated in Figure 15 (a). These regions represent the difference between highest and lowest points in corresponding QRS complex. In this research, the proposed MMD algorithm makes attempts to deal with such signal abnormalities and exploits such features as illustrated in Figure 15 (b) to accurately identify the location of the R wave (if present) from pre-processed raw ECG signals.

Therefore, the proposed system consists of five key steps including baseline correction (Section 5.1.2.1), Max-Min difference curve generation (Section 5.1.2.2), dynamic threshold computation (Section 5.1.2.3), R-Peak detection and Error correction (Sections 5.1.2.4 and 5.1.2.5). The flowchart of the overall system pipeline is illustrated in Figure 16. Each key step is discussed in detail in the following section.

As previously mentioned, before applying the proposed MMD algorithm, a sliding window average filtering strategy is first used to filter major spikes as they highly affect peak detection process. The following formula is used to define the corresponding operation.

$$y_c(x) = \begin{cases} \frac{1}{w}\sum_{i=1}^{w} y\left(x + i - \frac{w}{2}\right) & \text{If } \frac{w}{2} \le x \le size(y) - \frac{w}{2} \\ y(x) & Otherwise \end{cases} \quad (5.1)$$

where $w = round(fs * .02) + round(fs * .02)\%2$ (5.2)

where *y (x)* indicates the intensity of the raw ECG signal at sample *x* while y_c *(x)* represents the filtered output signal at sample *x*. *fs* indicates the sampling frequency of the signal and *w* represents the filtering window size, which is an even number close to the number of samples in 0.02 second.

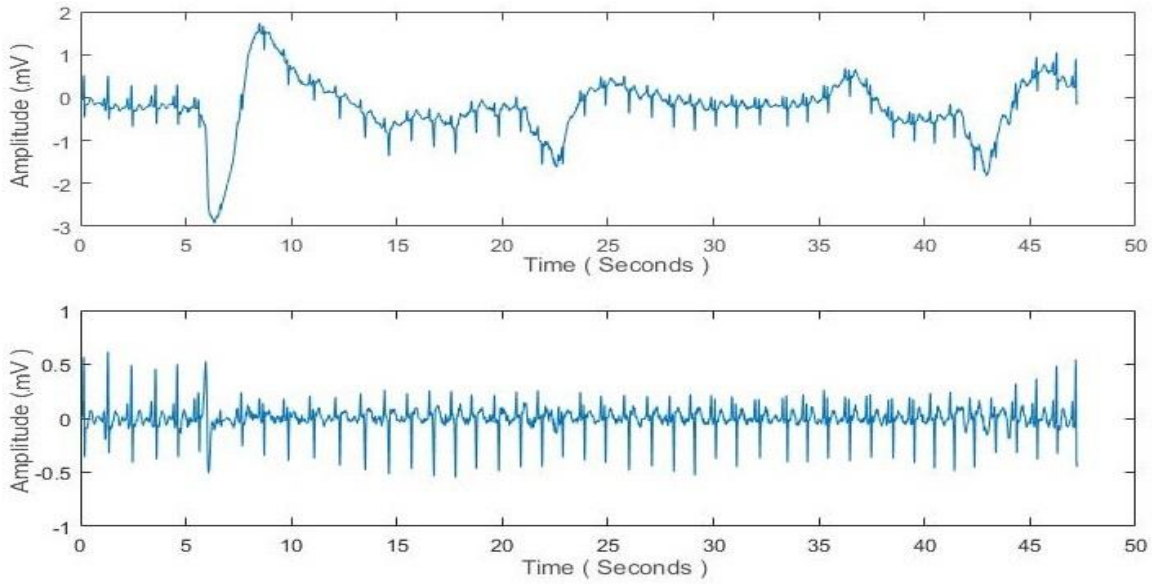


Figure 17 : ECG signal with baseline wander (top) and the corrected ECG signal (bottom)

5.1.2.1 Baseline Correction

Baseline wander occurs frequently in ECG readings particularly those derived from mobile ECG sensors. In this research, baseline wander is fixed by subtracting the local average of the signal from the raw signal. The following formulas are used for conducting the related operations.

$$y_b(x) = \begin{cases} y(x) - \frac{1}{w}\sum_{i=1}^{w} y\left(x + i - \frac{w}{2}\right) & If\ \frac{w}{2} \le x \le size(y) - \frac{w}{2} \\ y(x) & Otherwise \end{cases} \quad (5.3)$$

where $w = round(fs * 0.5) + round(fs * 0.5)\%2$ (5.4)

where $y(x)$ indicates the raw ECG signal and $y_b(x)$ denotes the baseline corrected signal. In this case, the window length w is an even number close to the number of samples in half a second. Figure 17 shows an example of ECG signal with baseline wander (top) and the corrected signal (bottom).

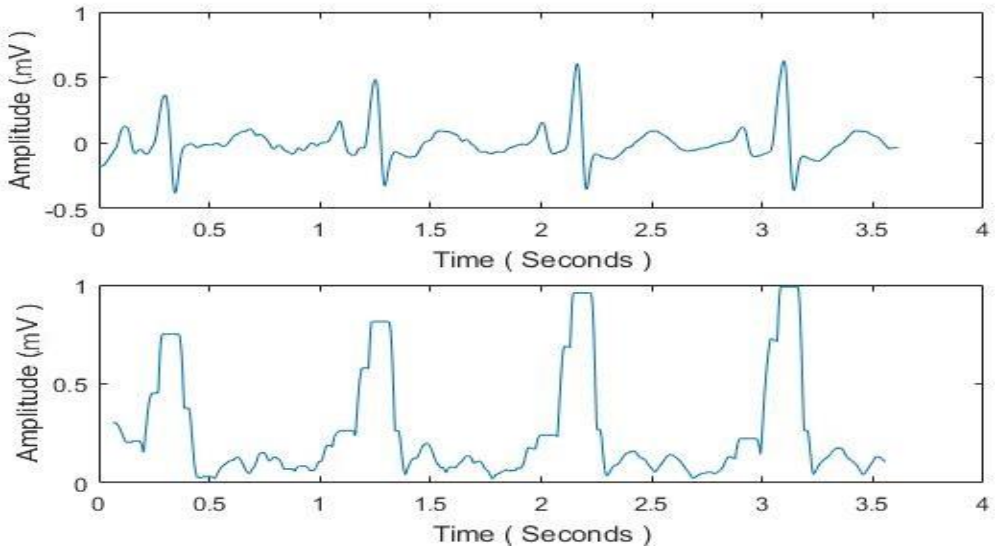


Figure 18 : An example ECG signal (top) and the corresponding MDC (bottom)

5.1.2.2 Max-Min Difference Curve Generation

In this section, Max-Min Difference Curve (MDC) generation is elaborated. The MDC represents the difference between the highest and lowest amplitudes in a window region. The following equation is used in MDC generation:

$$y_{mdc}(x) = Max\,(W(x)) - Min\,(W(x)) \tag{5.5}$$

$$W(x) = \begin{cases} \left\{y\left(x+1-\frac{w}{2}\right), \dots, y\left(x+w-\frac{w}{2}\right)\right\} & If\ \ \frac{w}{2} \le x \le size(y) - \frac{w}{2} \\ \{y(1), \dots, y(w)\} & Else\ if\ \ x < \frac{w}{2} \\ \{y(size(y)-w), \dots, y(size(y))\} & Otherwise \end{cases} \tag{5.6}$$

$$w = round(fs * 0.14) + round(fs * 0.14)\%2 \tag{5.7}$$

where $W(x)$ represents local window at position x. The window size w is an even number close to the number of samples in 0.14-second as most of the QRS complex fits in that length (Rangayyan 2002). Figure 18 shows an example of MDC along with the ECG signal. It is noticeable that all peaks in the MDC are positive and aligned to the QRS complexes of the corresponding ECG signal.

5.1.2.3 Dynamic Threshold Computation

In the next stage, the proposed dynamic thresholding for R-peak detection is discussed. As indicated earlier, not all peak values in MDC represent R-peaks in actual ECG signal and an ineffective threshold may lead to mistaking P or T waves as the R wave. In reality, the threshold for R-peak detection not only varies from time to time but also varies from signal to signal. In this research, in order to solve the above problem, two decision making parameters are employed to determine the subsequent occurrence of the QRS complex, i.e. the location threshold (*tLoc*) and the amplitude threshold (*tAmp*). This stage can be divided in the following 3 key steps; 1) Initialization, 2) Updating and 3) Time varying threshold selection. At start, the variables required for threshold calculation are initialized. In the next step, the threshold is updated in each iteration according to previous values. Eventually, it is finalized based on current positions and other factors. The detailed steps are described as follows.

- **Initialisation**

First, *tLoc* initialize as 0.2**fs*, considering the maximum heartbeat per minute can only reach up to 300 (or one heartbeat per 0.2 second). *tAmp* is initialized as *pf*max (W (x))* for the first iteration where *W* (*x*) is the search window and *pf* is the peak factor representing the minimum factor of previous peaks to be taken as the next threshold. Considering that the number of heart beats per minute can go down to no less than 24 (or one heartbeat per 2.5 seconds) and a QRS complex stays for 0.14 seconds only (Rangayyan 2002). The size *w* of the search window is initialized as follows:

$$w = round\ (fs * 2.64) + round\ (fs * 2.64)\ \%2 \qquad (5.8)$$

The start of the search window, *ws*, is also initialized below:

$$ws = rLoc + round(fs * 0.14) \qquad (5.9)$$

where *rLoc* is the location of the most recently found QRS complex.

- **Updating**

After the first *n* iterations, *tLoc* and *tAmp* are updated according to previous values of *mAmp* and *mLoc* as formulated below:

$$tAmp\ (x) = \frac{1}{n+1}\left(\sum_{i=1}^{n} mAmp(x-i) + mAmp(x)\right) \qquad (5.10)$$

$$tLoc(x) = \frac{1}{2}\left(\frac{1}{n-1}\left(mLoc(x-1) - mLoc(x-n)\right)\right) \qquad (5.11)$$

where *n* is the recall number representing the number of previous QRS properties to remember while *mAmp (x)* and *mLoc (x)* are the amplitudes and locations of the selected peak above the threshold on MDC at iteration *x*, respectively.

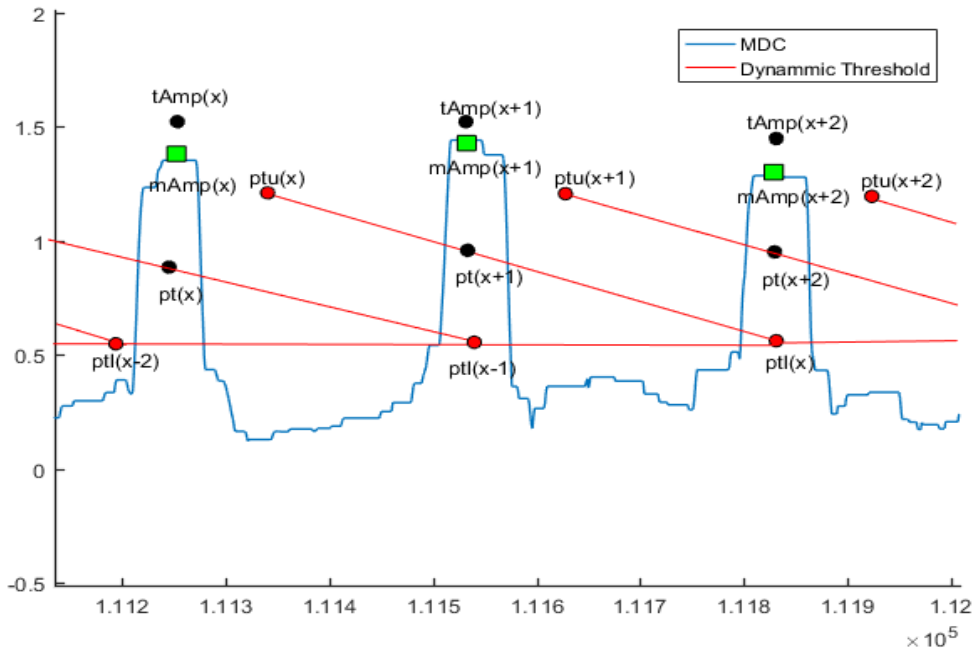


Figure 19 : Time varying dynamic threshold selection on MDC

- **Time varying threshold calculation**

Subsequently, depending on the position of the current peak, the value of the threshold is selected. Linear interpolation of values between an upper threshold (*ptu*) and a lower threshold (*ptl*) is used. After one MDC peak is selected as an R-peak, *ptu (x)* and *ptl (x)* are updated using the following equations:

$$ptu\ (x) = tAmp\ (x) * uf \qquad (5.12)$$

$$ptl\ (x) = tAmp\ (x) * lf \qquad (5.13)$$

The peak threshold for the current peak *pt (x)* is calculated as follows.

$$pt(x) = \begin{cases} ptu(x) + (ptl(x) - ptu(x)) * \dfrac{mLoc(x) - rLoc(last) - 0.14fs}{2tLoc(x) - rLoc(last) - 0.14fs} & If\ mLoc\ (x) - rLoc\ (last) < 2 * tLoc\ (x) \\ ptl(x) & Otherwise \end{cases} \qquad (5.14)$$

where *rLoc (last)* denotes the last detected R location. An example of how the proposed dynamic thresholding works is shown in Figure 19. Considering its x^{th} peak, *ptu (x)* and *ptl (x)* are calculated using Equations 13 and 14 as discussed earlier. The next search window starts

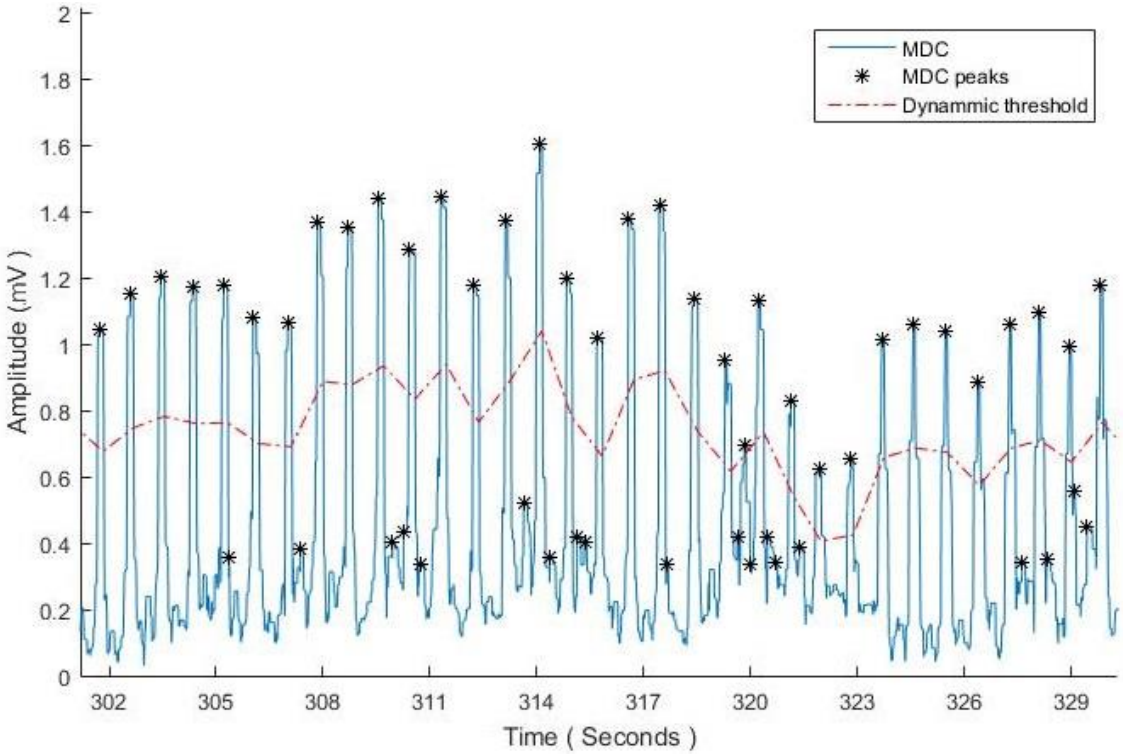


Figure 20 : MDC and dynamic thresholds.

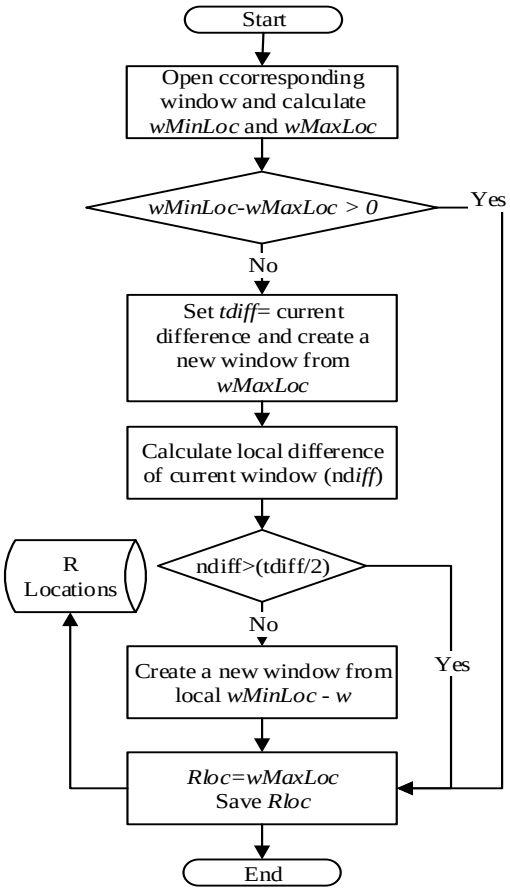
from the location of the *ptu* (*x*), which is *rLoc* (*last*) *+0.14*fs*. Depending on the distance of the last R location and the location of next peak on MDC, the peak threshold is varied from *ptu* (*x*) to *ptl* (*x*) and if it is more than twice of the average R-R distance or 2**tLoc* (*x*), the threshold at *ptl* (*x*) is fixed accordingly. When the next peak is selected, the values are updated and the same process is repeated. Figure 20 shows an example MDC along with varying thresholds. The main goal of this stage is to adapt the R-peak detection threshold not only to the signal but also to the localized window.

5.1.2.4 R-Peak Detection

In this stage, the MDC peak is either selected or discarded depending on the threshold value calculated above. In the peak detection stage, the algorithm creates a window of MDC starting from *ws* with a width of *w*. After that the following steps are performed.

- First, local maximum is detected which has a greater amplitude than *pt (x)* and stored as *mAmp (x)* along with location *mLoc (x).*
- If $mLoc(x) - mLoc(x-1) <= round(fs * 0.14)\ and\ mAmp(x) > mAmp(x-1)$, the previous R location is discarded continuing with the current iteration x.
- After the peak is selected on MDC, the actual signal is processed for the R-peak location.

Figure 21 shows the flowchart of finding R location from MDC peak, *mLoc*, and related pseudo-code is also given below.



wMaxLoc : Maximum amplitude location in the local window.
wMinLoc : Minimum amplitude location in the local window.
tdiff : Temporary difference between local minimum and maximum amplitudes.
ndiff : Difference between local minimum and maximum amplitudes.
w : Window size.
Rloc : Location of R

Figure 21 : Flowchart of R detection from the MDC peak

Algorithm 2: R-Peak Detection

Input:
mLoc //Location of peak found on MDC
Y (x) //The original ECG signal
Output:
rLoc //Location of the R wave if present, otherwise, location of the Q wave.

Begin

```
{
  window=W (mLoc) ;                    //Creates a window of size w
  start=mLoc;
  [wMin, wMinLoc]= min (window) ;      //Find value and location of minimum in the window
  [wMax, wMaxLoc]= max (window) ;      //Find value and location of maximum in the window
  If wMinLoc-wMaxLoc<0 then            //If the minimum before maximum, maximum might be R or T wave
        tdiff=wMax-wMin;               //Save the maximum-minimum difference
        new_window=W (start+wMaxLoc) ; //Create a new window from maximum location
```

```
        start= start+wMaxLoc;          // Update start of window
        [wMin, wMinLoc]= min (new_window) ;      //Find value and location of minimum in the new window
        [wMax, wMaxLoc]= max (new_window) ;      //Find value and location of maximum in the new window
        ndiff=wMax-wMin;               //Calculate the maximum-minimum difference
    If ndiff<tdiff/2 then                        //If new difference is less than half of previous difference
        new_window=W (start+wMinLoc-w)           //Create a new window before the current window
        start= start+wMinLoc-w                   //Update start of window
        [~, wMaxLoc]= max (new_window) ;         //Find location of the maximum in the new window
    Endif
Endif
rLoc = start+ wMaxLoc                //Save location of the maximum value
Output rLoc
}
End
```

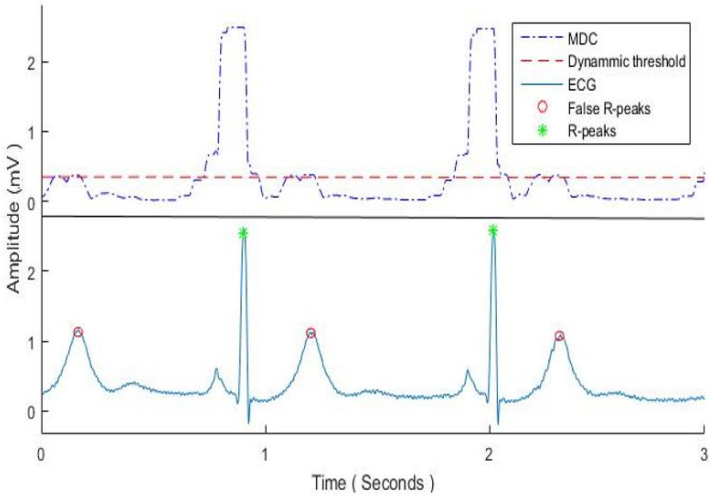


Figure 22 : MDC with insufficient upper threshold (top) and R-peak detection results shown on the ECG signal (bottom)

5.1.2.5 Error Correction

Finally, false positive errors are detected which are mainly caused by large T waves or an inefficient upper threshold value for R-peak detection. Figure 22 shows an example of unusually low thresholds on MDC leading to false positive R-peak detection. The diagram on the top shows MDC with a low threshold and the signal below is the corresponding ECG with false positive beat markers. Therefore, the following process is used to discard the falsely detected T waves and adjust the upper threshold accordingly. First, the last *n* differences of

detected R amplitude pairs are calculated. The resulting amplitudes, which are marked with double arrows as shown in Figure 23, are calculated using the following formula.

$$k\,(i) = rAmp\,(x\text{-}\,(i\text{-}1)\,*2)\,\text{-}R\,(x\text{-}\,(i\text{-}1)\,*2\text{-}1) \qquad (5.15)$$

The process is continued performing this calculation after at least *n**2 elements in *rAmp* is found.

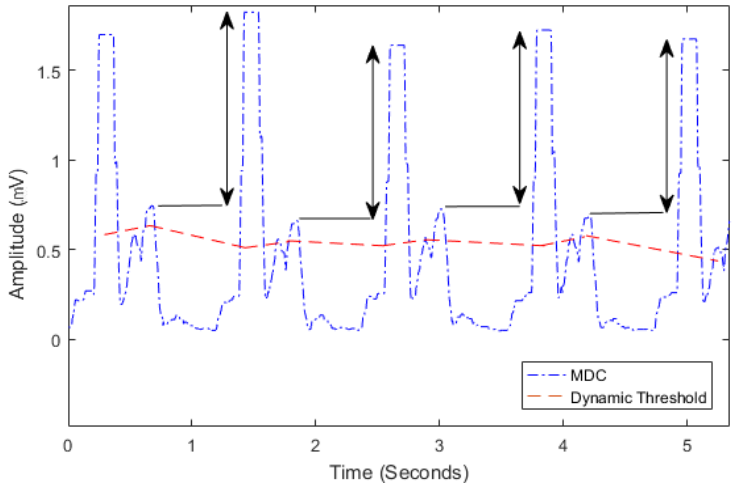


Figure 23 : Unusually low thresholds on MDC

The following inequalities are checked. If *absolute (k (i)) >0.25*rAmp (x- (i-1) *2) and absolute (k (i) -k (i-1)) <0.25*rAmp (x- (i-1) *2)* , hold for all *i*=[*1,2,.. (n-1)*], then the lower R location in each pair is discarded and *uf is* updated using the following formula:

$$uf = uf + \frac{1}{2}absolute\left(\frac{k(1)}{tAmp(x)}\right) \qquad (5.16)$$

In the experiments, only three pairs of R amplitudes are used for error correction (i.e. i=1, 2, and 3) and adjustment of upper R threshold factor (*uf*).

5.1.3 Experimental Results

In this section, evaluations on multiple diverse ECG signals have been conducted in order to prove the efficiency of the proposed system. First, the databases employed for testing are introduced, then the performance matrices comparing our work with other related research are discussed.

5.1.3.1 Experimental Data

Raw single lead ECG data samples which are extracted from five different databases in the PhysioNet (Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG,

Mietus JE, Moody GB, Peng C-K 2000) databank, are used for evaluation. These test databases are introduced below.

- The **MIT-BIH Arrhythmia Database (MITDB)** (Moody & Mark 2001) is made of 48 half-hour long 2-channel ambulatory ECG recordings taken from 47 subjects studied by the BIH Arrhythmia. The sampling frequency is 360 Hz. Two or more cardiologists manually annotated the records independently. This is one of the most widely used databases in ECG analysis.
- The **European ST-T Database (EDB)** (Taddei et al. 1992) contains 90 annotated 2-channel ECG signals taken from 79 subjects (70 men aged 30-84, 8 women aged 55-71 and information about the other person is missing). The signals are sampled at 250Hz per channel. The main objective of this database is to analyse ST-T wave changes.
- The **MIT-BIH ST Change Database (STDB)** (Albrecht 1983) includes 28 2-channel ECG recordings sampled at 360Hz. Signals are of varying length and recorded during exercise stress tests, which exhibit transient ST depression. The annotation files contain beat labels, however, the database does not include ST-T change annotations.
- The **St.-Petersburg Institute of Cardiological Technics 12-lead Arrhythmia Database (INCARTDB)** (Anon 2016) consists of 75 annotated thirty-minute long 12 lead recordings extracted from 32 Holter records. Signals are sampled at 257 Hz. Signals are annotated by an automatic algorithm and manually corrected subsequently.
- The **QT Database (QTDB)** (Laguna et al. 1997) consists of 105 fifteen-minute long recordings of two channel ECG signals with a wide variety of QRS waveforms which have been annotated manually to determine waveform boundaries. Channels have a sample rate of 250 Hz.

The above five databases contain one or multiple ECG leads. In the experiments, only a single channel (i.e. mostly Lead I or Lead II) is used from each ECG recording for system evaluation.

5.1.3.2 Performance Calculation

The QRS detection outcomes are referred to as True Positive (*TP),* False Positive (*FP*) and False Negative (*FN*). *TP* represents the total number of QRS complexes that are truly identified, while *FP* represents the total number of QRS complexes that are falsely marked even though they are not actually QRS complexes. *FN* represents the total number of QRS complexes that are not identified at all. Sensitivity (*SN*) and positive predictivity (*PP*) are used in this research for the evaluation of QRS detection. Equations *(5.17)* and *(5.18)* show the calculation of both *SN* and *PP*. Sensitivity represents the percentage of correctly detected beats whereas positive

predictivity represents the percentage of detected beats that are actually true. To simplify the results, TP, FP, and FN are converted to rates (TPR, FPR, and FNR) using Equations *(5.19)*, *(5.20)*, and *(5.21)* correspondingly.

$$SN = \frac{TP}{TP + FN} \times 100\% \quad (5.17)$$

$$PP = \frac{TP}{TP + FP} \times 100\% \quad (5.18)$$

$$TPR = \frac{TP}{TP + FN} \times 100\% \quad (5.19)$$

$$FPR = \frac{FP}{TP + FN} \times 100\% \quad (5.20)$$

$$FNR = \frac{FN}{TP + FN} \times 100\% \quad (5.21)$$

5.1.3.3 Evaluation Results

The first set of experiments are carried out on MIT-BIH Arrhythmia databases and the results of our algorithms are compared with a few existing methods. Appendix I shows signal-wise detailed results of multiple algorithms including Pan and Tompkins algorithm (1985), Wavelet Transform method by Li et al. (1995), the Difference Operation Method as described by Wang et al. (2008) and the JQRS method (Johnson et al. 2015; Behar et al. 2014). The window with a size of average QRS distance around R-peak is considered as true positive for MIT-BIH Arrhythmia database. This is because in MIT-BIH Arrhythmia database, the S wave is marked as QRS complex locations when the R wave is missing, while our algorithm tends to mark Q position as QRS location in the same situation.

As illustrated in Appendix I, the experimental results on MIT-BIH Arrhythmia database show comparable accuracy between our work and other related state-of-the-art algorithms reported in the literature. Our algorithm works very efficiently except for signals 105 and 207 where it yields some false positive results due to noise embedded in signals in the MIT-BIH Arrhythmia Database.

Table 5 : Test results of QRS complexes detection on multiple databases

Database	**Beats**	**MMD**					**JQRS (Johnson et al. 2015; Behar et al. 2014)**				
		TPR	**FPR**	**FNR**	**SN**	**PP**	**TPR**	**FPR**	**FNR**	**SN**	**PP**
MITDB	109809	99.65	0.34	0.35	99.65	99.66	98.86	0.38	0.84	99.16	99.62
EDB	790495	99.78	0.28	0.22	99.78	99.72	99.68	0.33	0.33	99.67	99.67
STDB	70755	99.92	0.24	0.08	99.92	99.76	99.46	0.54	0.58	99.42	99.46
INCRDB	174644	98.64	0.69	1.36	98.64	99.31	90.21	4.01	10.51	89.57	95.75
QTDB	86435	99.87	0.09	0.13	99.87	99.91	99.97	0.14	0.68	99.33	99.86
Overall	1232138	99.62	0.33	0.38	99.62	99.67	98.27	0.85	1.86	98.15	99.14

Table 6 : Performance comparison with related research

Publication	Methodologies	Beats	SN	PP
Pan & Tompkins (1985)	Bandpass filter + first derivative + squaring + moving average + multiple thresholds	116137	99.76	99.56
Li et al. (1995)	Wavelet transform + digital filter + singularity + thresholds	104182	98.89	99.94
Afonso et al. (1997)	Filter bank + thresholds	90909	99.59	99.56
Benitez et al. (2000)	First derivative + Hilbert transform + threshold	109257	99.13	99.31
Moraes et al. (2002)	Low pass filter + first derivative + modified spatial velocity + threshold	109481	99.69	99.88
Christov (2004)	Multiple moving averages + first derivative + thresholds	109494	99.76	99.81
Martínez et al. (2004)	Wavelet transform + Multiple thresholds + zero Crossing	109428	99.8	99.86
Chiarugi et al. (2007)	Bandpass filter + first Derivative + thresholds	109494	99.76	99.81
Arzeno et al. (2008)	First derivative + Hilbert transform + thresholds	109517	99.68	99.63
Yeh & Wang (2008)	Bandpass Filter + Difference Operation + thresholds	108517	99.86	99.95
Chouhan & Mehta (2008)	Digital filters + threshold	102654	99.55	99.49
Elgendi et al. (2008)	Digital filters + thresholds	44677	97.50	99.90
Ghaffari et al. (2008)	Continuous Wavelet transform + threshold	109837	99.91	99.72
Zheng et al. (2008)	Discrete Wavelet transform + Cubic Spline Interpolation + moving average + threshold	N/R	98.68	99.59
Elgendi (2013)	Bandpass filter + first derivative + squaring +thresholding using two moving averages	109985	99.78	99.87
Chouakri et al. (2011)	Wavelet transform + histogram + moving average + two thresholds	109488	98.68	97.24
Zidelmal et al. (2012)	Wavelet transform + coefficients multiplication + two thresholds	109494	99.64	99.82
Rodríguez et al. (2015)	Bandpass filter + first Derivative + Fast Fourier transform + adaptive threshold	44715	96.28	99.71
The proposed approach	Moving average filter + dynamic thresholding + MMC generation	1232138	99.62	99.67

Moreover, the proposed algorithm is compared with the most recent state-of-the-art JQRS method and tested both algorithms using the five databases. Overall, a set of 1229248 heartbeats from five different databases were used for evaluation. The detailed result comparison between our algorithm and the JQRS method is shown in Table 5. Our algorithm achieves an overall average sensitivity of 99.62% and an average positive predictivity of 99.67% when evaluated with the five databases and outperforms the JQRS method consistently. Especially, our results obtained using QTDB show a very low false positive rate, leading to an overall highest positive predictivity of 99.91%. Also, the performances of both our algorithm and the JQRS method on INCARTDB are comparatively lower since signals in INCARTDB are noisier and less enhanced in comparison to those in QTDB and STDB.

Finally, the proposed algorithm is also compared with multiple related developments over the years as described in Elgendi et al. (2014) in Table 6. Since these related developments (as shown in Table 6) employed different datasets and evaluation strategies, it shows a loose comparison with related research to estimate system efficiency. Empirical results indicate that the proposed system achieves comparable performances when tested with diverse databases with efficient computational efficiency.

Our experiments are conducted using Matlab on a PC with 2.5 GHz CPU. The average operation time for thirty-minute long ECG signal is about 3-5 seconds. The empirical results indicate that the proposed MMD algorithm not only obtains high accuracy for multi-datasets evaluation but also has efficient computational cost with a computational complexity of *O(n)* where *n* is the length (number of samples) of ECG signal.

5.1.3.4 Complexity Evaluation

Th computational efficiency of the proposed algorithm is also demonstrated in comparison to those of other related methods, i.e. JQRS (Behar et al. 2014; Johnson et al. 2015) and the work of Pan and Tompkins (1985), for R-peak detection. The computational cost comparison experiments have been conducted using MATLAB on a PC with 2.5 GHz CPU. Four databases, i.e. MITDB, EDB, STDB, and INCERTDB, are employed for comparison. To enable the possible comparison of computational complexity, JQRS (Behar et al. 2014; Johnson et al. 2015) and the work of Pan and Tompkins (1985) are also applied on a 5-second sliding window to each signal of the test databases, although the above conventional methods are not designed for sliding window based real-time analysis. The computational costs are illustrated in Table 7. Experimental results indicate that these conventional methods not only have higher computational costs compared to our proposed MMD algorithm but also produce a comparatively much higher number of false positive results for R-peak detection. This may be due to the fact that these related methods are not designed for sliding window based real-time analysis and may require the whole signal to obtain a good frequency resolution as opposed to employing a small sliding window based processing. On the contrary, the proposed MMD algorithm obtains optimal computational efficiency and shows great robustness in dealing with real-time QRS detection with impressive accuracy. Overall, it achieves a computational complexity of *O (n),* where *n* is the length of an ECG signal.

Table 7 : The average CPU time elapsed for R-peak detection for each method

Databases	JQRS (Behar et al. 2014; Johnson et al. 2015)	Pan and Tompkins (Pan & Tompkins 1985)	The proposed MMD algorithm
MITDB	19.1097	12.12012	10.23568
EDB	42.12465	34.51354	31.64323
STDB	14.32272	10.4369	9.578125
INCERTDB	10.79729	8.016667	7.815833

5.2 PQRST Extraction

Two simplified feature extraction algorithms are proposed which identify the locations and amplitudes of P, Q, R, S, T waves from lead II ECG signals. Eleven features are subsequently extracted based on the derived P, Q, R, S and T indicators to represent one-lead heartbeat signal more effectively. The first algorithm extracts the features without any provided R locations; and the second algorithm uses R location as the input for further processing.

5.2.1 Standalone P, Q, R, S, T Extraction

In this section, the initial ECG data processing for noise cancellation is discussed and the proposed feature extraction algorithm is elaborated for abnormal beat detection. In order to conduct effective noise reduction, dynamic moving window average is implemented to recover standard ECG signals. Then a wave detector algorithm is proposed to identify one-lead signals and extract several ECG features including shape, positions, amplitudes and duration times of P, Q, R, S, T waves. The proposed feature extraction algorithm also has efficient computational costs in order to be easily deployed to mobile platforms.

Algorithm 3: The Wave Detector Algorithm

Inputs:

1. A raw ECG signal (*buffer[]*)
2. Sample rate in Hz (*freq*)
3. Window size in seconds (*w*)

Outputs:

1. Location and amplitude of P wave peak P[] (locP, ampP)
2. Location and amplitude of Q wave peak Q[] (locQ, ampQ)
3. Location and amplitude of R wave peak R[] (locR, ampR)
4. Location and amplitude of S wave peak S[] (locS, ampS)
5. Location and amplitude of T wave peak T[] (locT, ampT)
6. Average signal B[] (ampB)
7. R_iR_{i-1} Interval RR[] (rr)

Begin

1. Initialize:

   Signal Pointer P=1, R Threshold tR=0.5 mV, RR Interval iRR=0.4 sec, QRS Threshold iQRS = 0.015 sec

2. Check if buffer has enough data

   If Size (buffer) -P >= w*freq

   Go to next step

   Else Exit.

3. Find baseline b of the signal by averaging the signal window B[i]

$$b = \left(\sum_{i=P}^{p-1+w*freq} buffer(i) \right) / (w* freq)$$

4. Find next R[i] (locR, ampR) by checking first peak more than b+tR and minimum peak distance freq*iRR from previous locR (for first peak the minimum distance is ignored). For the first iteration RR is considered 0.8*freq. Otherwise, RR[i] = R[i].locR-R[i-1].locR
5. To cancel noises, apply dynamic moving window average after R is recognized. In this process, each point of signal is recalculated as a simple arithmetic average value of neighboring signal points by spanning window size including left and right side of the signal. The averaging (smoothing) is done in such a way when window is closer to R the window size is smallest and when it goes away from R it gradually increases and reaches maximum up to 10. Window increment is adjusted based on RR interval. Window size is increased by one on each RR/ (2*8) distance from R.
6. Find corresponding Q[i] and S[i] by identifying local minima within iQRS*freq Range from left and right-hand side of current R[i] location from buffer. Although smoothing is done to detect locQ and locS the amplitudes (ampQ and ampS) are taken from the original signal.
7. Find P[i] from left side of Q[i] by checking local maxima above baseline and within RR/2 range.
8. Find T[i] similarly from the right side of S[i] by checking local maxima above baseline and within RR/2 range.
9. Modify signal pointer

   P=R[i].locR+ Round (RR[i]/2).

10. P = P + w*freq. **Go to step 2.**

End

The outputs of this algorithm include the features of P, Q, R, S, and T waves, which will be used to guide the recovery of eleven features for an effective representation of a one-lead heartbeat signal for classification. The classification algorithm for abnormal beat detection in detail is presented in the following.

5.2.2 Assisted Q, P, S, T Extraction

To further demonstrate the potential of the proposed MMD algorithm for abnormality detection in ECG signals, a simple algorithm is proposed which can detect locations of P, Q, S, and T

waves based on the detected R-peak locations (by MMD). Algorithm 4 shows the peak detection procedure for P, Q, S and T waves without compromising computational efficiency. The algorithm employs a local extrema search process to identify the locations of P, Q, S and T peaks for abnormality detection. The search process finds the local minima and maxima. The algorithm takes a raw ECG signal and the R-peak location as the inputs, and provides the estimated peak locations of P, Q, S, and T waves as the outputs. Although more complex processing for peak detection of P, Q, S and T waves is available (Lin et al. 2010; Rahimpour & Mohammadzadeh 2016), some of them tend to be computationally expensive. The detection of each wave (e.g. P or T wave) also constitutes a research topic on its own (Messaoud et al. 2009; Elgendi et al. 2015). In this research, a comparatively lightweight peak detection algorithm is used for detection of P, Q, S and T waves. The results indicate that it is efficient in accurately recovering the locations of P, Q, S and T waves, without distracting the focus of the research on the proposed MMD algorithm for QRS detection too much. Moreover, only the signals with accurate QRS detection are employed in this step for peak detection of P, Q, S, and T waves and for subsequent abnormality analysis.

Algorithm 4: R Assisted feature extraction

Input:
Y(x) //The original ECG signal
locR //Location of R
qrsT //QRS distance threshold
pT //P distance threshold
tT //T distance threshold
Output:
locP //Location of the P
locQ //Location of the Q
locS //Location of the S
locT //Location of the T
ampP //Amplitude of the P
ampQ //Amplitude of the Q
ampS //Amplitude of the S
ampT //Amplitude of the T
Begin
//Search the left-hand side of the R location for the minimum value within the half QRS distance threshold range
locQ = location of minimum in *Y(locR - Round(qrsT/2): locR);*
//Search the right-hand side of the R location for the minimum value within the half QRS distance threshold range
locS = location of minimum in *Y(locR: locR + Round(qrsT/2));*
//Search the left-hand side of the Q location for the maximum value within the P distance threshold range
locP = location of maximum in *Y(locQ - pT: locQ);*

```
  //Search the right-hand side of the S location for the maximum value within the T distance threshold range
  locT = location of maximum in Y(locS: tT);
  ampQ = Y(locQ);
  ampS = Y(locS);
  ampP = Y(locP);
  ampT = Y(locT);
End
```

5.2.3 Experimental Results

In this section, the conducted experiments are described and the results are shown. The feature extraction algorithms determine the locations of P, Q, R, S and T waves, however, the databases used in the study only provide locations of QRS complex. This makes it hard to access the feature extraction algorithms without manually labelled boundaries. As the acquired P, Q, R, S and T wave locations from feature extraction algorithms are used for abnormality detection, the classification results are provided to indicate the efficiency of the extracted features. The results from the standalone algorithm (i.e. the positions of P, Q, R, S and T locations of each heartbeat) are used in the next chapter for further processing. In this chapter, only the results calculated using assisted feature extraction algorithm are used. The R values used are extracted by the MMD algorithm.

In this study, an artificial neural network (ANN) and ensemble classifiers are employed for heart abnormality detection based on the features extracted by assisted feature extraction method from single lead raw ECG signals. Neural network is employed in this research because of its generalization ability and robustness for dealing with any unseen test data. Especially, the NN has a great capability of adjusting itself to large variations of input samples for classification which fits the requirement of this research very well. A large number of data samples from online ECG databases are also used for the training of NN. Afterwards, the trained network can be used on any new input samples for abnormality detection.

5.2.3.1 Databases used

Raw ECG data samples used from signals provided by three databases from PhysioNet (Goldberger et al. 2000) databank for evaluation. These databases are, European ST-T Database, QT Database, MIT-BIH Arrhythmia Database. All of the databases mentioned above are freely available online. A more detailed introduction of these databases is provided in Section 5.1.3.1.

The training and testing of our system is described in the following. Only one channel of ECG data is used from the databases discussed in this section to create training and testing data. For

the large signals only first one million of the data sample is used. If the signal is small, the whole signal is processed. The heartbeats are categorized into two types: 'Normal' and 'Abnormal' regardless of abnormality types at this stage for this research. Table 8 shows the number of 'Normal' and 'Abnormal' heartbeats taken from each database for training and testing purposes. The test database contains only a subset of beats extracted by our proposed algorithm from PhysioNet databases.

Table 8 : The number of normal and abnormal ECG beats in the employed databases.

Database	No. of normal beats	No. of abnormal beats
European ST-T	133846	10009
QT	40271	6051
MIT-BIH arrhythmia	18783	7396
Total	202909	23545

5.2.3.2 Assessment method

The classification outcomes are referred to as *TP*, *FN*, *TN*, and N_2. *TP* represents the total number of abnormal ECG beats that are truly identified as abnormal ECG beats, while *FN* represents the total number of abnormal ECG beats that are falsely identified as normal ECG beats. *TN* represents the total number of normal ECG beats that are correctly identified as being normal, and *FP* represents the total number of normal ECG beats that are falsely identified as abnormal. True positive rate (TPR), false positive rate (FPR) and overall accuracy (OA) are used as measures in this research for the evaluation of the heartbeat abnormality detection. Equations (5.19) and (5.20) show the calculation of TPR and FPR while overall accuracy is calculated using Equation (5.22).

$$OA = \frac{TN + TP}{TN + TP + FP + FN} \times 100\% \qquad (5.22)$$

5.2.3.3 Assessment Methods

Two simple assessment methods are used for checking accuracy on abnormality (heartbeat abnormality) detection. Neural network is selected as first method and Adaboost M1 with decision trees as the second. Both the methods are discussed further in the following.

- **Abnormal beat detection based on ANN**

Artificial neural network is a popular computing model mainly used for solving classification problems. In the experiments, an ANN was employed with 11 inputs, 5 hidden layers and two output classes. Features obtained from the proposed feature extraction algorithm along with

annotations from database are used to train the network by Backpropagation algorithm. After the training is done, the network is capable of classifying normal or abnormal heartbeats when those sixteen extracted features (discussed further in Chapter 6) are used as inputs of the network. The evaluation results for NN are discussed in the next section.

- **Abnormal beat detection based on ensemble classifier**

An ensemble classifier combines a set of trained weak learner models to form a strong classifier. It produces ensemble results for any test data by aggregating predictions from its weak learners. In the experiments, decision tree (DT) was taken as the base weak classifier along with AdaBoost for the construction of the ensemble classifier. The outputs of the DT classifiers are combined into a weighted sum which represents the final output of the boosted classifier. Adaptive boosting is built in a way that subsequent weak learners are tweaked in favour of those instances which are misclassified by the previous classifiers.

5.2.3.4 Experimental Results

In the experiments, 60% heartbeats are taken randomly for training of neural network and ensemble classifier. The remaining 40% are used for testing purpose. Table 9 shows the evaluation results for the Artificial Neural Network and the ensemble classifier. True positive and false positive rates vary for different databases. The false positive rate is very low while using ensemble classifier, however, the true positive rate is almost same for both the classifiers. Overall, the ensemble classifier shows slightly better performance than neural network classifier as shown in Table 9.

Table 9 : Detection results for normal and abnormal ECG beats for the testing databases.

Test databases	Artificial neural network classifier			Ensemble classifier		
	TPR	FPR	OA	TPR	FPR	OA
European ST-T	99.3	0.24	99.7	97.1	0.15	99.7
QT	95.2	0.53	99.0	96.0	0.42	98.9
MIT-BIH Arrhythmia	94.5	0.94	97.5	97.7	0.29	99.6
Average	96.33	0.57	98.73	96.93	0.29	99.40

5.3 Conclusion

In this chapter, a novel method, i.e. the MMD algorithm for QRS detection from single lead (mostly Lead II) ECG signals is proposed. It combines baseline correction, Max-Min difference curve generation, dynamic threshold computation, QRS detection and error correction. The main contribution of the proposed MMD algorithm focuses on building a lightweight real-time

QRS detection scheme without compromising detection accuracy. The proposed approach shows great efficiency in dealing with QRS detection from diverse cross-domain ECG signals with efficient computational complexity. Evaluated with five well-known databases, the proposed MMD algorithm achieves impressive performances in comparison to related research for R-peak detection using both normal and abnormal ECG signals. Furthermore, integrated with feature extraction and a neural network classifier, the proposed MMD algorithm has also been further extended for abnormal/normal heartbeat detection. The empirical results indicate the efficiency and superiority of the proposed algorithm to aid abnormality detection. However, it shows performance degradation when dealing with noisy signals, especially with high amplitude spikes, which motivates future directions for development.

Two simple feature extraction algorithms are also proposed in this chapter. The assisted method (as discussed in Section 5.2) is used for evaluation which uses R values obtained by MMD algorithm. On the other hand, the less accurate standalone P, Q, R, S and T detection method is used in the following chapter to show how performance (in terms of accuracy) can be improved using feature enhancement. After using the feature extraction method, the results are saved in the database for further feature generation or enhancement which will be elaborated in detail in the following chapter. The MMD algorithm was published in "Computer Methods and Programs in Biomedicine" (Pandit et al. 2017) and some of the content were also presented in "SKIMA 2014" (Pandit et al. 2014).

Chapter 6 Feature Enhancement

Feature enhancement is an optional part of ECG processing. The enhancement methods are generally integrated with the corresponding extraction algorithm rather than reporting separately. In this chapter, it is illustrated that a simple feature enhancement scheme can be helpful in improving overall accuracy of abnormality detection in heartbeats.

Previously, the proposed R-peak detection algorithm along with P, Q, S, T extraction algorithms obtains the locations of P, Q, R, S, and T waves. In this study, two experiments are conducted, the first experiments using eleven constructed features (which are similar to standard medical features) and the second experiment using sixteen features (previously used eleven features and added enhanced/generated features). Both set of features are generated from the same initial information (raw ECG signal and locations of P, Q, R, S, T waves). The goal of this experiment is to demonstrate how performance eventually differs upon using different sets of generated features.

6.1 Features

Features in ECG signal contains information (meaningful data) that can be used in classification stage for abnormality detection. Standard medical features (containing medically valuable information) are explained in Feature Extraction section of Chapter 2. Most of the feature extraction algorithms intend to find the locations of P, Q, R, S, T waves; then, recalculate the standard medical features. Sometimes it is also infeasible to calculate boundary

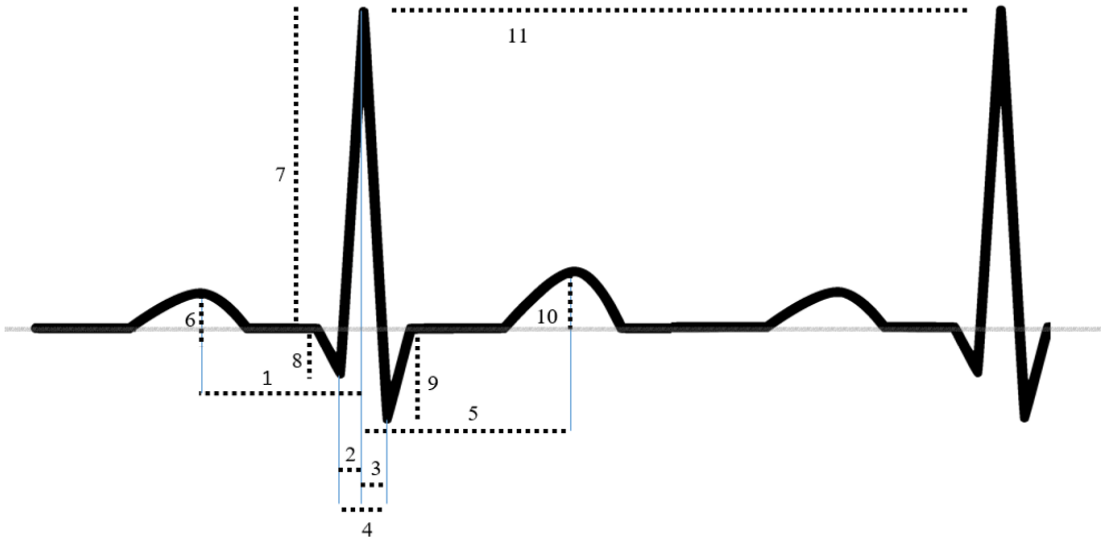


Figure 24 : Eleven non-standard features

of each wave as they seem to overlap each other. In the next section, commonly used features and reconstructed features used in this research are discussed.

6.1.1 Eleven Non-Standard Features

Initially, approximate locations of P, Q, R, S, T waves are extracted for each heartbeat from the raw ECG signal as the outputs of the feature extraction algorithms. Only locations of those waves in raw ECG signals cannot be used directly for classification (except for R location which can be used for heartbeat variability analysis). So, some features are constructed which are very close to standard features, however, without using boundary analysis for sub waves. The features are shown in Figure 24 and further described in Table 10. The horizontal grey line represents the baseline of the signal.

Table 10 : Eleven features used in the first study.

No.	**Expression**	**Description**
1	(locR-locP) /*rr*	Normalized P-peak R-peak distance.
2	(locR-locQ) /*rr*	Normalized Q-peak R-peak distance.
3	(locS-locR) /*rr*	Normalized R-peak S-peak distance.
4	(locS-locQ) /*rr*	Normalized Q-peak S-peak distance.
5	(locT-locR) /*rr*	Normalized R-peak T-peak distance.
6	|ampP– ampB |/ampR	Normalized P-peak amplitude.
7	ampR–ampB	R-peak amplitude.
8	|ampB–ampQ|/ampR	Normalized Q-peak amplitude.
9	|ampB–ampS|/ampR	Normalized S-peak amplitude.
10	|ampT–ampB|/ampR	Normalized T-peak amplitude.
11	*rr*	Distance of R-peak from previous one.

where locP, locQ, locR, locS, locT are locations of P, Q, R, S, T waves respectively and the amplitudes are marked as ampP, ampQ, ampR, ampS, ampT. AmpB and *rr* are the baseline amplitude of the signal and distance between current R-peak and previous R-peak locations respectively. These values are the outputs of the previously proposed feature extraction

algorithm, which are marked on top of an ideal ECG signal in Figure 24 using dotted lines. The marked numbers in Figure 24 represent the corresponding feature indexes in Table 10.

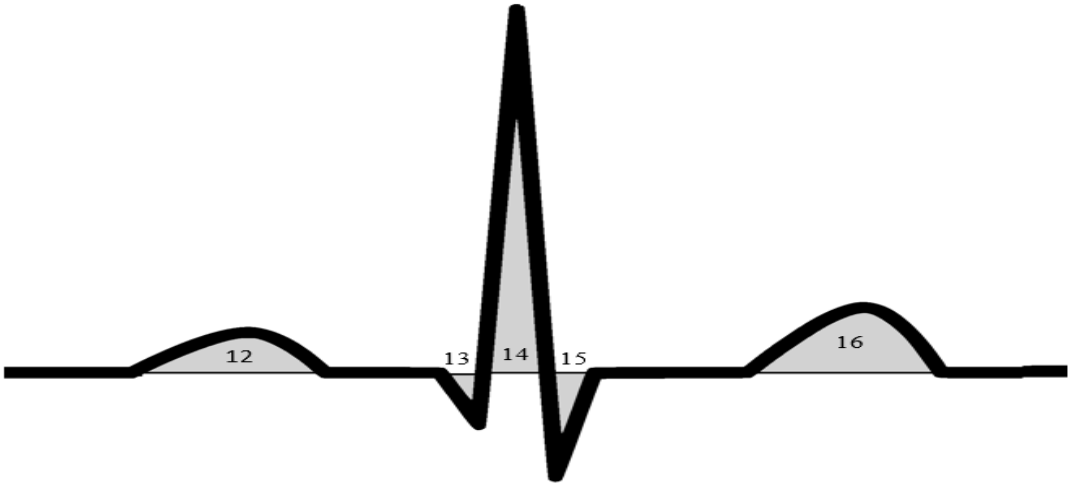


Figure 25 : Generated additional features plotted on an ideal ECG signal

6.1.2 Generated Features

In this section, some improvements are proposed over the previously used feature set by generating additional features and extending the previous ones. While using the previously mentioned eleven non-standard features, some of medical information is discarded regarding the duration and shape of P, Q, R, S, T waves for abnormality detection. Table 11 shows the new set of features that is used in the corresponding experiments. The new feature set is derived in the following way. For normalization, the distance based features (i.e. features 1-5) are divided by the sampling frequency of the signal (*fs*) instead of using R-R distances (*rr*) as in Table 10. This change ensures abnormal R-R distance does not affect the normal distance based features. Similarly, amplitudes (i.e. features 6-10) are not divided by R-peak values to discard any effect of abnormal R-peak over other features. The calculation of peak amplitude features (i.e. features 6-11) is also no longer divided by the current R as in the previous work shown in Table 10. A set of 5 new features (i.e. features 12-16) are also generated, which represent the valuable durations and shape information of P, Q, R, S, T waves. They are visually marked as grey regions as shown in Figure 25 in corresponding to the feature index shown in Table 11. As mentioned earlier, these new features represent P, Q, R, S, T area in the ECG signal conveying valuable information about duration and shape of the corresponding waves. They are calculated using the expressions as described in Table 11 where *kp*, *kq*, *kr*, *ks* and *kt*

indicating the half size of corresponding waves in the number of samples and *Y (i)* is the raw ECG signal. For example, *kp* is calculated as follows *kp = round ((fs*0.11) /2)* considering average P wave lasts for 0.11 seconds. Similarly, kq, *kr*, *ks* and *kt* are calculated based on average duration of the corresponding waves.

Table 11 : The new feature set used in the second study.

No.	Expression	Description
1	(locR – locP) /fs	Normalized P-peak R-peak distance.
2	(locR – locQ) /fs	Normalized Q-peak R-peak distance.
3	(locS – locR) /fs	Normalized R-peak S-peak distance.
4	(locS – locQ) /fs	Normalized Q-peak S-peak distance.
5	(locT – locR) /fs	Normalized R-peak T-peak distance.
6	ampP – ampB	P-peak amplitude (where *ampB* referring to the baseline amplitude of the signal).
7	ampR – ampB	R-peak amplitude.
8	ampB – ampQ	Q-peak amplitude.
9	ampT – ampB	T-peak amplitude.
10	ampB – ampS	S-peak amplitude.
11	(locR-$locR_{previous}$) /fs	Normalized distance between two adjacent R-peaks
12	$\sum_{i=locP-kp}^{locP+kp} \lvert Y(i) - ampB(i) \rvert / fs$	Area of P wave (where *kp* is the half of the average P wave duration and *Y* (*i*) is the i^{th} sample of the original signal).
13	$\sum_{i=locQ-kq}^{locQ+kq} \lvert ampB(i) - Y(i) \rvert / fs$	Area of Q wave (where *kq* is the half of the average Q wave duration).
14	$\sum_{i=locR-kr}^{locR+kr} \lvert Y(i) - ampB(i) \rvert / fs$	Area of R wave (where *kr* is the half of the average R wave duration).
15	$\sum_{i=locS-ks}^{locS+ks} \lvert ampB(i) - Y(i) \rvert / fs$	Area of S wave (where *ks* is the half of the average S wave duration).
16	$\sum_{i=locT-kt}^{locT+kt} \lvert Y(i) - ampB(i) \rvert / fs$	Area of T wave (where *kt* is the half of the average T wave duration).

6.2 Abnormality Detection Experiments

In this section, two sets of experiments are conducted using two separate set of features as previously described. The same datasets and feature extraction algorithms are used for determining P, Q, R, S, and T locations, however, different feature sets are generated. The rest of the section is divided into three parts: 1) Databases, 2) Experimental Setup and 3) Experimental results.

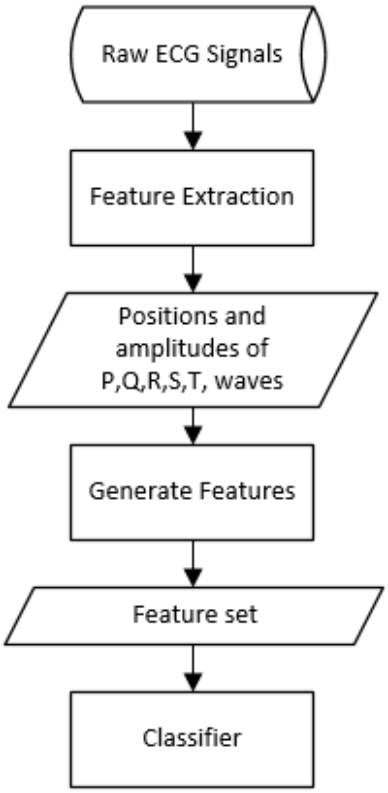


Figure 26 : Experiment sequence for abnormality detection

6.2.1 Databases

In the next experiments, raw single lead ECG signals extracted from PhysioNet (Goldberger et al. 2000) databank, were used for training and testing. Databases including MIT-BIH Arrhythmia Database (Moody & Mark 2001), QT Database (Laguna et al. 1997), and European ST-T Database (Taddei et al. 1992) are used for evaluation of the previous algorithm along with new feature enhancements. In this experiment, much larger (than the dataset in Table 8) and unbalanced dataset is tested for thorough evaluation. Table 12 shows the number of normal and abnormal heartbeats extracted from multiple databases for testing.

Table 12 : The number of normal and abnormal ECG beats extracted from ECG databases for system testing

Database	No. of normal beats	No. of abnormal beats
MIT-BIH arrhythmia	46170	5809
QT	80707	6288
European ST-T	784633	5932
Total	911510	18029

6.2.2 Experimental Setup

The sequence of operations of the whole experiment is shown in Figure 26. In the first stage, the previous feature extraction algorithm is tested on the raw single lead ECG signals, i.e. the algorithm generates approximate positions and the peak values of the P, Q, R, S and T waves. In the next stage, a set of 11 features (per heartbeat) is extracted from each signal using the formulas as discussed in Table 10. Subsequently, the duration and shape related information of the corresponding waves were added to achieve additional five features (features 12-16) using the formulas in Table 11. In the final stage, multiple classification techniques were used for detecting abnormality in the heartbeats.

6.2.3 Experimental Results

Multiple classification algorithms were used for testing the previous and newly derived feature sets (as described in Table 10 and Table 11) using the above-mentioned test databases. The experiments are conducted using 10-fold cross validation to evaluate the performance of multiple classifiers. As the number of normal examples in most of the databases is very high compared to the number of abnormal class samples, only a portion of normal ECG data was taken randomly in order to balance the training samples for both of the cases for the training of the classifiers. A total of 10 classifiers were used for abnormality classification, which includes K-Nearest Neighbour (KNN), Naive Bayesian (N-Bayes) , Decision Tree (D-Tree) , Random Forest (R-Forest) , Artificial Neural Network (ANN) , Support Vector Machine (SVM) , and ensembles such as AdaboostM1, Voting, Bagging, and Stacking.

We have used Sensitivity (*SN*) and Positive Predictivity (*PP*) to evaluate the classifiers' performance. *TP*, *FP*, *TN*, *FN*, are defined similarly as in Section 5.2.3.2. *SN* and *PP* calculation is defined in Equation *(5.17)* and *(5.18)*

In AdaBoost and Bagging ensemble methods, decision tree is used as the base classifier of the ensembles. In this research, Matlab is used to implement feature extraction algorithm whereas the classifiers are built with the support of Weka (Smith & Frank 2016) and Rapidminer (RapidMiner 2016).

The performances of diverse classifiers using the previous feature set are shown in Table 13. Sensitivity (SN) represents the percentage of abnormal heartbeats that are identified by the classifier and Positive predictivity (PP) indicates the percentage of detected abnormal heartbeats which are truly abnormal. KNN and the Ensemble classifiers show very high average performance for both SN and PP while N-Bays and SVM have poor average SN scores.

Table 13 : Performances of classifiers using 11 features

Classifiers		MITDB	EDB	QTDB	Average
KNN	**SN**	88.78	89.44	95.5	91.24
	PP	90.67	92.84	96.99	93.50
N-Bayes	**SN**	28.23	66.1	45.36	46.56
	PP	62.6	75.23	71.23	69.69
D-Tree	**SN**	93.9	78.18	73.4	81.83
	PP	84.75	82.91	90.53	86.06
R-Forest	**SN**	28.24	55.88	41.98	42.03
	PP	94.07	91.16	96.24	93.82
ANN	**SN**	77.02	87.36	84.44	82.94
	PP	86.17	93.87	88.73	89.59
SVM	**SN**	59.41	52.51	57.7	56.54
	PP	81.91	75.21	81.91	79.68
Adaboost	**SN**	89.44	90.00	89.44	89.63
	PP	94.75	95.3	90.75	93.60
Bagging	**SN**	89.32	89.74	89.32	89.46
	PP	90.77	93.42	90.77	91.65

Table 14 : Performances of classifiers using 16 features

Classifiers		MITDB	EDB	QTDB	Average
KNN	**SN**	91.40	87.23	97.87	92.17
	PP	93.15	86.37	98.08	92.53
N-Bayes	**SN**	31.06	64.80	55.21	50.36
	PP	78.08	79.86	75.22	77.72
D-Tree	**SN**	93.00	78.32	76.24	82.52
	PP	84.75	82.88	90.68	86.10
R-Forest	**SN**	53.36	55.38	62.97	57.24
	PP	90.47	94.18	97.47	94.04
ANN	**SN**	90.78	88.42	89.36	89.52
	PP	95.44	94.39	95.03	94.95
SVM	**SN**	63.34	54.58	61.79	59.90
	PP	85.16	88.26	62.95	78.79
Adaboost	**SN**	96.43	91.02	92.34	93.26
	PP	92.05	86.58	93.58	90.74
Bagging	**SN**	89.24	92.72	96.24	92.73
	PP	91.85	89.52	96.27	92.55

The experimental results using a modified and extended feature set are shown in Table 14. There is a very small deterioration of KNN and SVM performance, however, in comparison to results shown in Table 13, mostly the results for the new feature (set as illustrated in Table 14) indicate impressive performance improvement. This is because the new feature set is more

robust and carries extra information about wave durations and shapes while reducing RR distance based errors.

6.3 Conclusion

In this research, a feature construction method is proposed showing improvement over our previous research along with performance evaluation using various classification methods. In this study, a strategy to extract valuable information about the duration and shape of ECG waves is proposed without the requirement of finding the actual boundary of the waves. The proposed modified feature set along with some generated extra features contributes toward higher accuracy of abnormality detection which is reflected in the experimental results. In future work, these extracted features will be further tested to prove their efficiency for heart disease detection from ECG signals. Outlier and novel class detection will also be explored to ensure the efficiency in dealing with challenging noisy ECG based disease diagnosis. Some materials of this research were presented in "Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)" (Pandit et al. 2016).

Chapter 7 Heartbeat Classification

Heartbeat classification is the last stage of automated ECG processing. In this stage, the dataset represented by the acquired features is used to find out if there are any abnormalities present in the corresponding ECG signal. Classification algorithms are used for this purpose. These algorithms take the features as inputs and conduct abnormality detection. Classification is conducted using supervised machine learning strategies. So, before they are able to predict any abnormality, they need to be trained by an example dataset containing examples of normal and abnormal cases. Although single standard classifiers can be used for this purpose, the meta classifiers are preferred as they tend to improve performance in terms of precision, recall and the overall accuracy. In this chapter, a hybrid classifier is proposed which is constructed using multiple ensemble classifiers. The result of the hybrid classifier is compared with other conventional meta classifiers. The hybrid classifier results in higher accuracy than that of the conventional meta classifiers. Additionally, a novel class detector is also presented in this chapter. The conducted experiments show the usefulness of the novel class detector when novel classes are present in the test datasets.

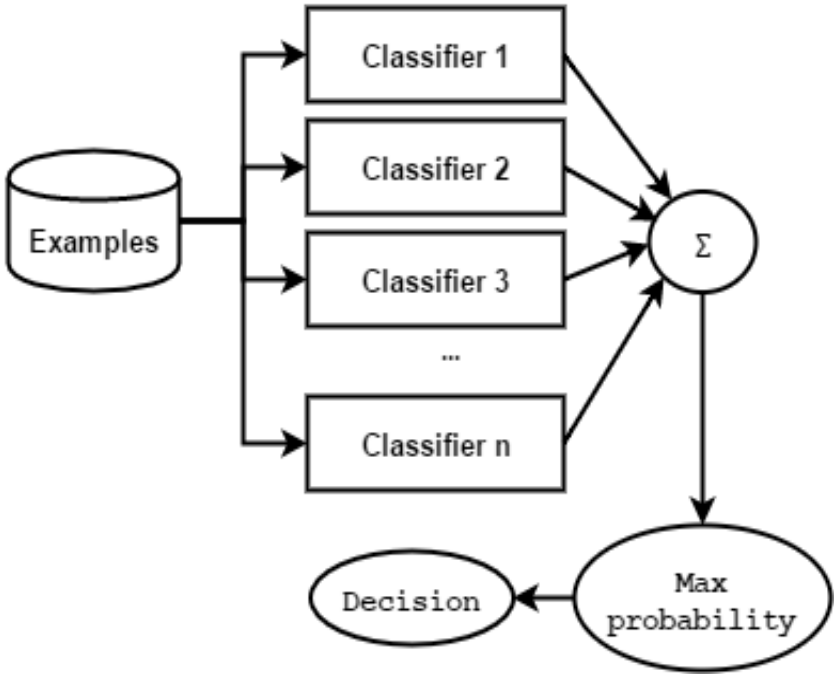


Figure 27 : The General Structure of an Ensemble Classifier

7.1 The Proposed Hybrid Classifier

In this section, the development of hybrid classifier using multiple ensembles is discussed in detail. One or multiple learning algorithms can be used collaboratively to obtain better predictive performance than those of single classification models. Figure 27 shows the basic structure of a general ensemble classifier. Typically, an ensemble classifier consists of set of diverse base classifiers/learners (e.g. Classifier 1, Classifier 2, …, Classifier *n*) and a decision strategy (which combines decisions from the base learners/classifiers) to determine the final prediction results. In the following section, the proposed methodologies, the implementation strategy, and the experiments are explained.

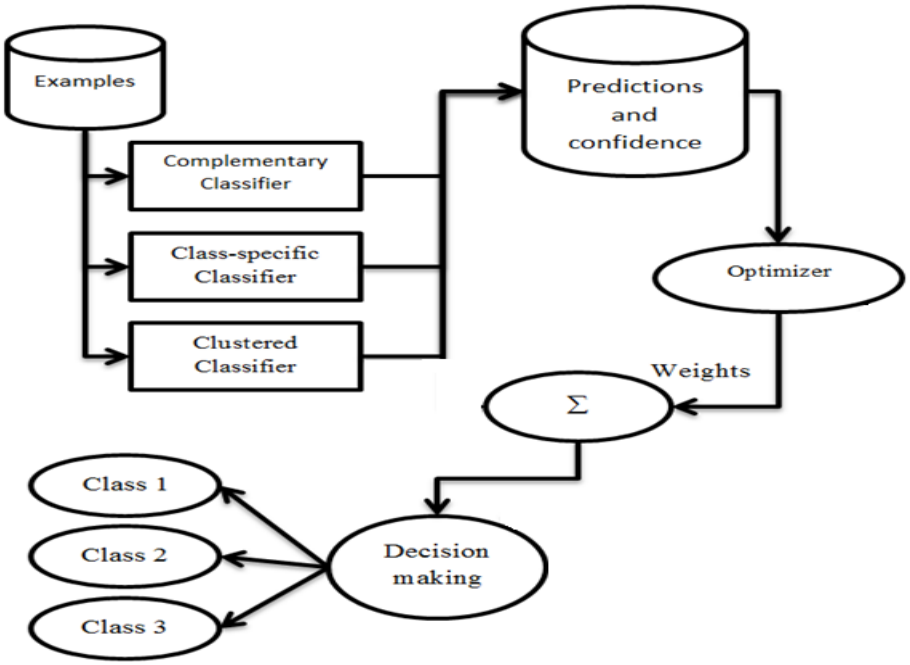


Figure 28 : Structure of the proposed hybrid classifier

7.1.1 Methodology

The structure of the proposed hybrid ensemble system is discussed in this section along with the introduction of multiple algorithms used for building the hybrid classifier. Figure 28 shows the overall diagram of the proposed hybrid classifier. The proposed hybrid classifier includes three ensembles, i.e. Complementary Classifier, Class-specific Classifier, and Clustered Classifier. Each of them is focused on a specific perspective of classification tasks. After

training, these three classifiers can independently classify an example test dataset. Classification outputs along with confidence scores for each ensemble model are passed onto the next stage to draw the final decision. Then, the weight of each classifier is computed using an optimization algorithm in the next stage to achieve enhanced classification accuracy. After that, the identified optimal weights of the three base ensemble models are used in the weighted majority voting to reach to the final decision. The class with the highest probability is selected as the final decision output.

7.1.1.1 Training of Base Learners

There are three ensembles used as the base learners while creating the hybrid classifier. The results of each ensemble contribute towards the final decision through weighted majority voting. Each of the three ensembles is made of multiple base learners which can be replaced by any type of multi-class classifier (e.g. Decision Tree, Artificial Neural network, SVM). These three ensemble classifiers gather knowledge about training data in different ways which eventually create variance in the way they are able to make decisions. There are three ensembles namely Complementary Classifier, Class-specific Classifier, and Cluster-specific classifier. The construction and training of the abovementioned ensemble classifiers are discussed below.

- **Complementary Ensemble Classifier**

The main idea behind creating the complementary ensemble classifier is to create a set of classifiers where the mistakes of each classifier can be neutralized by another classifier so that the overall result can be as correct as possible. The complementary ensemble classifier combines multiple base learners where the next learner complements the misclassified dataset by the previous learner. The first classifier is trained normally. The next classifier is trained using the same dataset examples but with higher weights on the examples which were mistaken by the previous classifier and the process is repeated for the building of subsequent base models. The training algorithm is discussed in the following.

Algorithm 5: Creating and Training the Complementary Ensemble Classifier

---

Input: D //Training dataset containing d no. of elements

n //Number of intended base learners

Output: Trained Complementary Classifier Base learners

Method:

1: Initialize weights

$W=\{w_1, w_2, \ldots, w_d\}=\{1,1,1..\text{ d times }\}*1/d$

2: **for** j=1 to n **do**

3: **if** j==1 **then**

4: Train a new base learner L_j using randomly taken 1/n portion of dataset and corresponding weights from w.

5: **else**

6: Train a new base learner L_j using m and corresponding weights and if no. of elements in m< d/n randomly use elements from training set till total number reaches d/n.

7: **end if**

8: Test the base classifier using whole dataset and mark subset m⊆D of example which are incorrectly classified.

9: W={1,1,1.. d times } //Reset Weights

10: **for each** element in m **do**

11: Update corresponding weight $w_i = w_i * 2$. (for i^{th} element)

12: **end for**

13: W=$W / \sum_1^d w_i$. //Update weights

14: **end for**

Each of the base learners can be made of the same or different classification model (e.g. ANN, SVM, Decision Tree etc.). The above algorithm describes the method of creating the complementary ensemble classifier. The variations of the base learners are created in this ensemble by training those using different intersecting subsets of the training dataset with higher weights. The subset of examples with higher weight used to train each base classifier includes all examples wrongly classified by a preceding classifier. When the ensemble classifier is fully trained, a dataset can be tested by passing it to all the base learners and recording all confidence outputs for prediction classes. A majority voting strategy is used to combine all base learners' outputs for each class. The class with the highest confidence is selected as the output class along with its confidence score.

- **Class-specific Ensemble Classifier**

The class-specific ensemble classifier creates a set of base learners, each of which is trained using boosted instances (with higher weight) belonging to a specific class. Therefore, each base learner is trained to gain knowledge of a specific class and is used to determine whether a test example belongs to that particular class or not (working as binary classifier). The weights of training set are boosted till there is no false negative left (in validation) for the corresponding class of a specific classifier. In the end, each base learner is able to identify a specific class with 100% recall on that class. Algorithm 6 describes its training process.

Algorithm 6. Creating and Training Class-Specific Ensemble Classifier

Input: D //Training dataset containing d no. of elements
T //No. of boosting iterations

Output: Trained Class-specific Classifier Base learners (L)

Method:

1: Initialize weights
C= number of classes in training data
W={1,1,1.. d times } /d //Initial Weights
2: **for** i=1 to C **do**
3: Create a dataset D' from D where if the examples belong to i^{th} class, mark class level true or else marks it false.
4: **for** t=1 to T **do**
5: Train a new base learner L_i using D' dataset and weights W.
6: Test the base classifier using whole dataset and mark misclassified examples m
7: W={1,1,1.. d times } //Reset Weights
8: **for each** element in m **do**
9: Update corresponding weight $w_j = w_j * 2$. (for j^{th} element)
10: **end for**
11: $W = W / \sum_1^d w_j$. //Update weights
12: **end for**
13: **end for**

When the training process is complete, each of the base classifiers accumulates the knowledge of one specific class. In the testing phase, majority voting is not possible (like in Figure 27) as the base classifiers do not provide prediction probabilities for each class. There could be a few cases while testing this ensemble classifier, i.e. 1) The output of i^{th} classifier is true and all the other classifiers' outputs are false. In this case, the example is marked as it belongs to class *i*. 2) The outputs of multiple classifiers are true. In that case the classifier with highest confidence score (say classifier *i*) is selected and the example is classified as that particular class (class *i*). 3) All the base classifiers' outputs are false. In this case, we select the classifier (say classifier *i*) which has the lowest confidence score and mark the example with that class label (class *i*).

- **Clustered Ensemble Classifier**

The next ensemble classifier is developed to find dissimilarity among similar examples. As a result, the final accuracy is increased in discriminating different classes. First, the cluster-specific ensemble classifier creates clusters of the training dataset. Then, it creates a single

classifier (base learner) for each cluster and trains the classifier using samples from the corresponding cluster. Clustering is an unsupervised algorithm that groups instances based on their similarities. In an ideal situation, each group of data should belong to a single class, however, for majority of the cases, this does not happen. As a result, multiple class instances might be present in each resulting cluster. Each of the base classifiers of the ensemble classifier is trained using each cluster data. The detailed algorithm is discussed as follows:

Algorithm 7. Creating and Training Class-specific Ensemble Classifier

---

Input: D //Training dataset containing d no. of elements
n //No. of base learners

Output: Trained Cluster-specific Classifier Base learners (L)

Method:

1: Initialize weights
W={w_1, w_2, …, w_d }={1,1,1.. d times }*1/d
2: Use clustering algorithm to create n number of clusters
3: **for** i=1 to n **do**
4 for each element in D do
5: **if** current example belong to cluster i
Update corresponding weight w_j= 2 *w_j. (for j^{th} element)
6: **end if**
7: W=$W / \sum_1^d w_j$. //Update weights
8: Train a new base learner L_i using D dataset and weights W.
9: W={1,1,1.. d times } //Reset Weights
10: **end for**

The testing process is similar to that of the complementary ensemble classifier. The outputs of each base learner are averaged after running the examples through all of them and the class with the maximum probability is selected as the final output.

7.1.1.2 Optimizing weights of base learners

The three ensemble classifiers individually have their strengths and weakness in terms of classification performance, however, the collective intelligence proves effective as all three classifiers possess different points of view of the same classification problem. Although, it is possible to use a simple majority voting process to determine the final classification result, a weighted majority voting based on optimization yields better performance results. In this research, an evolutionary optimization algorithm, i.e. the scattering and repulsive firefly

algorithm (SRFA), is proposed to optimize the weights of the three ensemble classifiers. SRFA is a modified firefly algorithm to mitigate the premature convergence of the original FA and is used to optimize the weight for each ensemble model to generate the final weighted majority voting result. The details of SRFA and the comparison with other state-of-the-art search algorithms are discussed in the next chapter.

7.1.2 Experiments

In this section, the experiments are described. Multiple well-known datasets from UCI machine learning repository are used for experimentation along with the ECG dataset with enhanced features as discussed in Chapter 6. This section describes the datasets, experimental setup and evaluation results in detail.

7.1.2.1 Datasets

Seven datasets are used for evaluation of the proposed hybrid classifier. Table 15 summarises datasets extracted from UCI machine learning repository (Bache & Lichman 2013) that were used for evaluation. They include well-known databases such as Iris, Glass Identification, Soybean large, Image Segmentation and KDD Cup databases. Another dataset is the ECG abnormality dataset extracted from the MIT-BIH Arrhythmia database. The raw lead II signals from this dataset are processed using our own feature extraction algorithm mentioned in previous chapters.

Table 15 : Dataset description

Dataset	Attributes	Attribute types	Instances	Classes
Iris	4	Real	150	3
Glass	10	Real	214	6
Wine	13	Integer, Real	178	3
Soybean	35	Categorical	307	15
Segmentation	19	Real	2310	7
KDD99	42	Categorical, Integer	4000000	23
ECG	16	Real	109809	14

- **Iris dataset**

The Iris dataset contains 3 classes and 50 instances for each class. Only one class is linearly separable from other classes.

- **Glass identification dataset**

The glass identification dataset has a total of 214 instances. It has 10 attributes containing real values. The total number of classes is 6.

- **Wine dataset**

The wine dataset contains 13 attributes, mix of integer and real numbers. It has a total of 178 instances and 3 classes.

- **Soybean large dataset**

The soybean large dataset contains a total of 35 attributes. This medium size data set has 307 instances which are divided into 15 different classes.

- **Image Segmentation dataset**

The image segmentation dataset has 2310 instances, 19 real attributes and 7 classes.

- **KDD Cup dataset**

The Knowledge Discovery and Data Mining 1999 competition dataset contains simulated network attack data. It contains a total of 41 attributes of integer and categorical data and a total of 4000000 instances. However, 6000 examples (1000 from 6 major classes) are used in this experiment.

- **ECG dataset**

The ECG dataset is derived from the experiments, which are conducted and discussed in previous chapters. It includes 16 features extracted from raw ECG signals taken from MIT-BIH arrhythmia database. It has 14 different classes with each representing a single type of heartbeat. There are 109809 sample heartbeats present in the database, while the majority of them belong to the normal class type. In our experiment, a subset of 1000 examples is taken from 6 major classes (including the normal class).

7.1.2.2 Experimental Setup

The experiments are conducted using the combination of Rapidminer (RapidMiner 2016) and Matlab (MathWorks 2016) software. Rapidminer software is used because most of the widely used and well-known algorithms are readily available in the software, while Matlab is used for implementation of our algorithm for evaluation.

7.1.2.3 Results

In the first experiment, the hybrid classifier is tested along with widely used meta ensemble classifiers with Decision Tree as the base learner. The meta classifiers include Adaboost, Bagging and Stacking. The results of the 10-fold cross validation of the abovementioned algorithms on each database are shown in Table 16. The experimental results illustrate that the hybrid classifier outperforms other classifiers in most cases for the detection of normal classes.

Table 16 : Performance Comparison between different ensembles

Databases	Overall Accuracy (OA)				
	DT	Boost	Bag	Stack	Hybrid
Iris	94.67	94.67	94.67	94.67	96.00
Glass	97.23	97.68	97.64	98.16	97.23
Wine	91.57	92.70	94.46	91.18	97.22
Soybean	81.43	85.33	90.53	83.76	90.51
Segmentation	88.02	90.88	93.29	88.75	94.02
KDD Cup	99.60	99.82	99.82	99.76	99.89
ECG database	95.23	96.84	96.11	97.67	98.18

7.2 Novel Class Detection

Novel class detection can be achieved by further modifying the proposed hybrid classifier. In this section, the structure of the novel class detector is explained along with the discussion on how it is used for instance classification and novel class detection. The novel class detector system is illustrated in Figure 29. The same ensemble classifiers employed in the hybrid classifier (see Figure 28) are used for novel class detection as well. Additionally, a novel class detection function is developed by determining novel class probability thresholds using the proposed SRFA optimization algorithm. That is, the weights along with novel class thresholds for each class are calculated using the SRFA algorithm based on confidence values predicted by the previous classifiers. In the next stage, the final decision confidence is compared with a class specific novelty threshold. If it is less than the threshold the example is classified as a

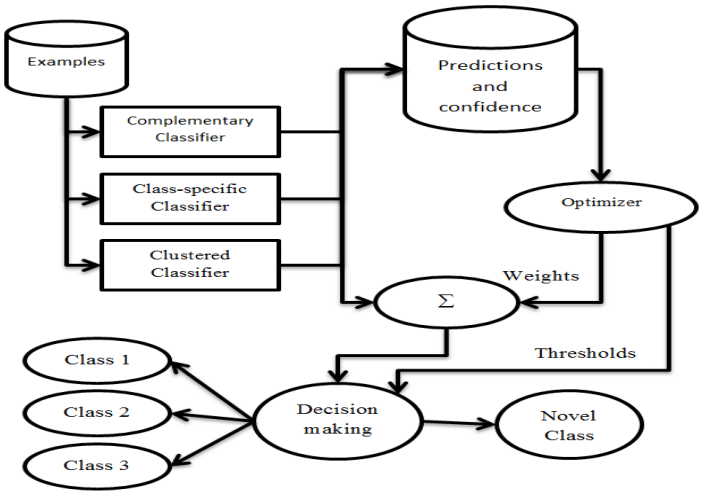


Figure 29 : Structure of Novel Class detector

novel class. The rest of this section is divided into two parts: Methodology and Experimental Results. The methodology section describes the techniques and algorithms used in the development of novel class detector and the Experimental Result section elaborates on experiments setup, databases used and classification results.

7.2.1 Methodology

The novel class detector is an extension of the proposed hybrid classifier which uses two extra algorithms to achieve novel class detection. The setup process of the novel class detector is shown in Figure 30 as a flowchart. In the first stage, the base learners are trained. Next, the weights of the base classifiers are optimized to enhance performance. After that, the predictions from base learners are saved along with the identified optimized weight for each base learner. In the final stage, the novel class threshold is calculated and finalized using the SRFA algorithm. Each stage is further elaborated in the following.

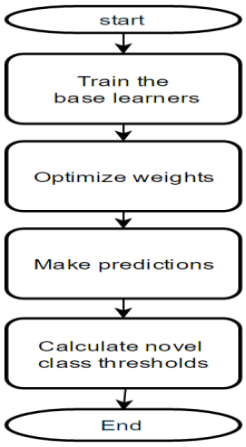


Figure 30 : Flowchart of novel class detector setup

7.2.1.1 Training of the base learners

In the novel class detector, the training processes of the base learners are also the same as those used for the hybrid classifier, which are described in Algorithm 5, Algorithm 6, and, Algorithm 7.

7.2.1.2 Optimizing weights of base learners

The base learner weights are optimized similarly as previously conducted in the hybrid classifier by using SRFA (see Chapter 8).

7.2.1.3 Decision making and confidence calculation

Decision making is performed using weighted majority voting. After the optimal weights are identified by SRFA, the weight for a particular classifier is multiplied with the output decision

confidence of the same ensemble classifier. Suppose, the output of an ensemble classifier n, has a confidence (probability) value of k for class c, i.e. $P_c^n = k$ and the finalized optimized weight for classifier n is w_n. The final output value for class c is set as $\sum_{n=1}^{n=3} P_c^n * w_n$. The final decision is the class that has the maximum value of $\sum_{n=1}^{n=3} P_c^n * w_n$.

The confidence score is calculated by checking the squared differences in prediction probabilities between selected class and the other classes after multiplying the weights. The confidence calculation for the final decision is performed based on the following algorithm:

Algorithm 8. Calculating the confidence score

Input: Confidences $C_i = \{ c_1, c_2, .. c_n\}$ for n classifiers.
Weight $W_i = \{ w_1, w_2, .. w_n\}$ for each classifier

Output: Final confidence C_f

Method:

//Multiplying weights

1: **for** i=1: n **do**
2: **for** j=1: n **do**
3: $p_{ij}= c_j*w_i$.
4: **end for**
5: **end for**
// Find maximum probability for each class
6: **for** j=1 to m **do**
7: $avg_j=\sum_{i=1}^{n} p_{ij}$
8: **end for**
9: Find max element pmax and its index pidx in avg
10: dc=0.
10: **for** i=1: m **do**
11: **if** i!=pidx **then**
12: dc=dc+ (pmax- avg_i) ^2/pmax.
13: **else**
14: **continue**
15: **end if**
16: **end for**
17: C_f=dc/ (m-1) ;

7.2.1.4 Optimizing novel class confidence thresholds

In this phase, the decision is made as to whether a predicted class belongs to a normal class or a newly arrived novel class, based on its confidence calculated by Algorithm 8. For every class in the training dataset, a threshold value (x_i for i^{th} class) is calculated. If the final decision indicates that a test sample belongs to a specific class, however the confidence goes below the novel class threshold, the example is declared to be a novel class instance.

The optimization of the novel class confidence threshold is conducted in a similar manner as that of the weight optimization using SRFA, although different fitness functions are employed. The goal of the optimization algorithm is to minimize the confidence threshold values of the novel class while maintaining the accuracy as high as possible. Suppose, thresholds for all classes are $\{x_1, x_2, x_3, \ldots x_m\}$. So, the fitness function is simply calculated as follows:

$$f(X) = \frac{Min\ (accuracy + e, 1)}{\sum_1^m x_i}$$

where *e* is the allowed error which is ignored in fitness calculation. It is a compromise between the classification accuracy for normal classes and the possibility for novel class prediction. An increasing allowed error (*e*) improves the chances for novel class detection, reducing the normal class detection accuracy (resulting in more false positive novel classes). Whereas, keeping *e* closer to 0 will yield the highest normal class detection accuracy and less chances of finding novel classes. This in turn results in an increased true negative. The goal of the objective function is to maximize accuracy while reducing novel class threshold for each classifier.

The optimization algorithm helps the novel class detection threshold to adapt itself accordingly to maximize accuracy of the final classifier. In our experiments, *e* is initialized as 0.2 based on trial and error, which generates the highest overall accuracy as described in the experimental results section.

7.2.2 Experiments

A set of experiments is conducted on multiple databases to demonstrate the performances of each test method along with the increase of the number of novel classes. First, the test databases are introduced with some modifications applied for novel class detection. Next, the experimental setup is explained.

7.2.2.1 Databases

In this section, the modification of the existing databases (described in 7.1.2.1) is described for the creation of multiple datasets for training and test purposes. One or multiple normal class

labels are not included in the training stage and used as novel classes for each database. The details are discussed below:

- **Iris and Wine dataset**

Two new datasets were created using Iris dataset and Wine datasets respectively. A subset of samples from two classes is used for training for both datasets while the remaining class (the third class) is used as the novel class at the test stage in this experiment.

- **Glass Identification, Soybean large and Image Segmentation dataset**

Four datasets are created from each of Glass Identification, Soybean large and Image Segmentation datasets. The four databases contain an increasing number of classes that are employed as novel classes (ranging from 1 to 4). For example, in the first dataset created from Glass Identification database, one class is used as the novel class while others are kept as the normal classes. In another dataset, i.e. the Image Segmentation dataset, there are four classes used as the novel classes. Therefore, the total number of resulting datasets is 12 (4 datasets from each database).

- **KDD Cup dataset**

From Knowledge Discovery and Data Mining 1999 competition dataset, this research randomly selects six major class examples (1000 examples from each class) for experimentation. Four sub-datasets are created from the extracted dataset similarly as those described earlier (with the increased number of novel classes from 1 to 4).

- **ECG dataset**

There are 6000 examples taken for experiments from the ECG dataset. This is because the majority of examples belong to the normal heartbeat type. Most of the abnormal examples are imbalanced with some of them containing a very small number of samples. Four major classes (i.e. normal and three abnormal cases) are chosen for experiments.

7.2.2.2 Experimental Setup

As mentioned earlier, all of the experiments are conducted using a combination of both Rapidminer and Matlab software. Similarly, Rapidminer was used for testing standard baseline methods on different databases using visual scripting (as shown in Figure 31) while Matlab was used to develop and test the proposed algorithms on the corresponding dataset. The experiment for the baseline meta classifiers is performed in the following steps.

1. Missing values in the database are replaced with average values (mostly occurring values for non-numeric values).

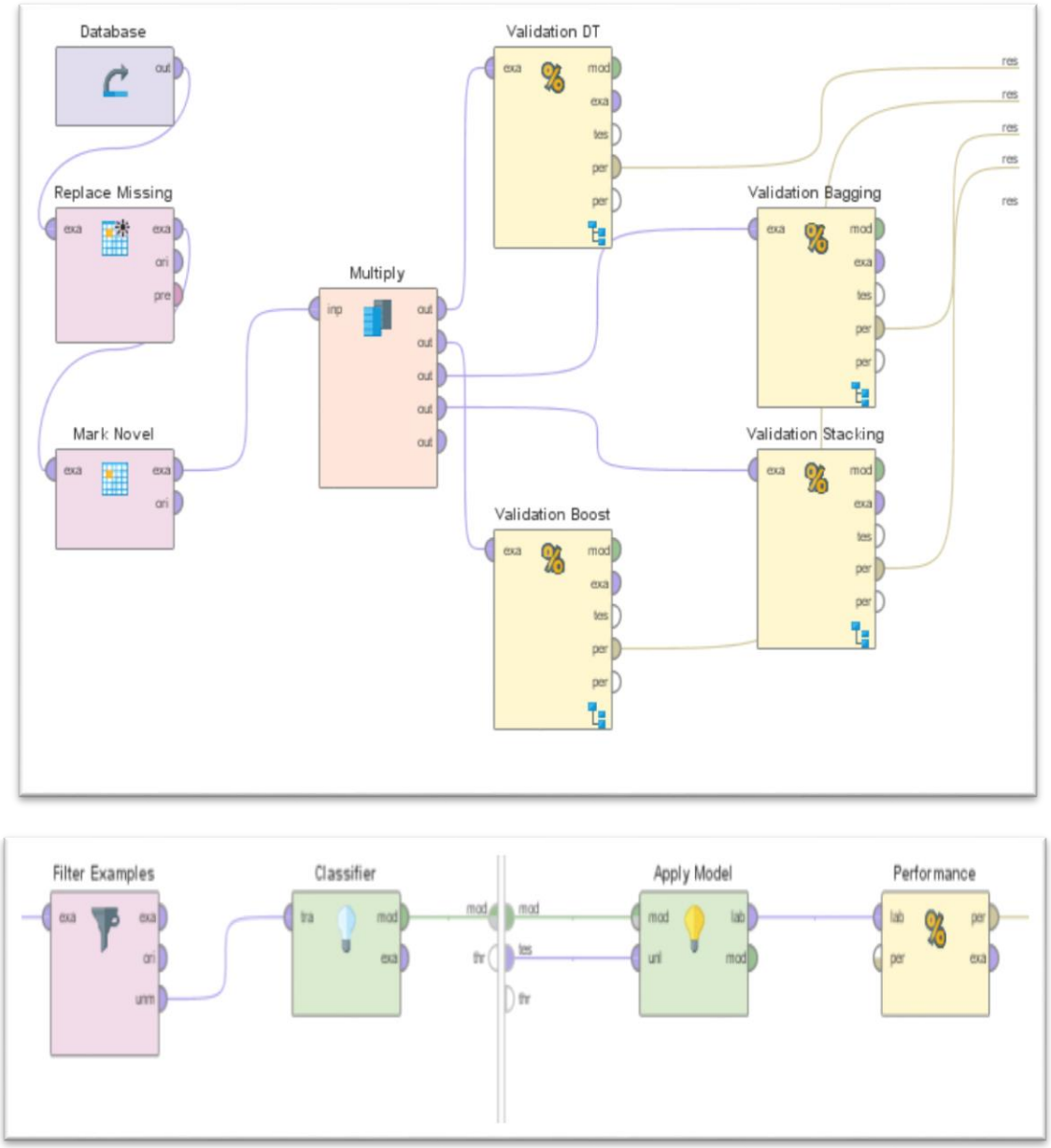


Figure 31 : Rapidminer visual scripting

2. One or multiple classes are marked as novel classes depending on the experiment.
3. The classifier is trained using samples from subsets of classes (as the novel class examples are filtered out)
4. The trained classifiers are validated using normal and novel classes together.
5. Tested with the corresponding test set(s), classification performance is calculated.

The entire process is shown in Figure 31 using visual scripting of Rapidminer. On the top, database preparation is shown and at the bottom, the validation process is provided for performance calculation. As mentioned earlier, the evaluation experiments of the proposed

novel class detector are conducted using Matlab. The steps of the experiment are described and shown in Figure 30.

7.2.2.3 Results

All instances of one or multiple classes are expelled randomly in the training phase. Those class instances are regarded as novel classes and in the test phase, novel class instances along with other normal class instances are used for testing. Table 17 defines the terms used for the performance calculation.

Table 17 : Terms and Definitions

Term	Meaning
TP	No. of novel class instances that are accurately classified.
FP	No. of normal classes that are falsely classified as novel classes.
TN	No. of normal classes that are accurately classified as normal classes.
FN	No. of novel class instances that are falsely classified as normal classes.

TPR, FPR and OA are defined using Equation (5.17-5.19). TNR and FNR are defined in Equations (5.20 and 5.21). TNR and FNR are different because, in these experiments, TN exists unlike in R-Peak detection experiments (discussed in Chapter 5) .

$$\text{TNR} = \frac{TN}{TN + FN} * 100\% \qquad (7.1)$$

$$\text{FNR} = \frac{FN}{TN + FN} * 100\% \qquad (7.2)$$

The classification experiment is conducted on all test datasets individually using 10-fold cross validation. In 10-fold cross-validation, the original sample is partitioned randomly into ten equal size subsamples. Out of ten subsamples, one subsample is retained as the validation data for testing the model, and the remaining nine subsamples are used as training data. The results are shown in depth in Table 18.

Table 18 : Classification performances on multiple databases in the presence of novel classes

Normal Classes	Novel Classes	Other Classifiers (OA)				The Proposed Hybrid Classifier				
		DT	ANN	SVM	AB	FPR	TNR	FNR	TPR	OA
Iris										
3	0	93.29	95.95	94.63	94.67					95.40
2	1	66.67	66.67	66.67	66.67	4.67	95.34	0.00	100.00	96.42
Wine Database										

3	0	94.92	94.90	89.18	94.38					94.50
2	1	69.10	69.12	70.75	69.18	0.00	100.00	23.73	76.28	92.14
Glass										
6	0	97.16	94.81	99.52	98.14					96.25
5	1	85.02	84.09	84.61	84.07	4.98	95.03	0.00	100.00	95.33
4	2	80.74	81.23	80.82	80.78	4.07	95.94	5.89	94.12	95.80
3	3	75.28	75.71	75.71	75.37	3.28	96.73	2.18	97.83	97.20
2	4	68.23	67.84	67.71	68.71	0.00	100.00	5.89	94.12	97.67
Soybean Database										
6	0	90.54	90.17	71.26	88.61					89.92
14	1	80.63	83.37	83.07	81.33	7.23	92.78	7.50	92.50	92.74
13	2	67.49	70.28	69.22	69.20	1.26	98.75	46.00	54.00	91.01
12	3	66.43	67.14	67.49	66.76	2.52	97.49	15.56	84.45	93.43
11	4	60.57	61.22	61.26	61.22	1.44	98.57	41.25	58.75	87.55
Segmentation Database										
6	0	92.79	88.04	69.88	90.91					91.42
6	1	72.29	75.64	77.02	76.05	0.00	100.00	10.35	89.66	98.57
5	2	63.21	62.67	59.21	61.76	0.00	100.00	8.48	91.53	97.61
4	3	52.74	51.19	54.00	53.69	0.00	100.00	5.62	94.39	97.61
3	4	40.60	41.24	41.63	41.71	0.00	100.00	7.57	92.44	95.70
KDD Cup Database										
6	0	99.62	88.82	96.62	99.84					99.61
5	1	79.89	79.98	79.93	79.93	4.98	95.03	0.00	100.00	95.33
4	2	59.93	59.93	59.93	59.93	4.07	95.94	5.89	94.12	95.80
3	3	39.91	39.93	39.93	39.93	3.28	96.73	2.18	97.83	97.20
2	4	19.98	19.99	20.00	20.00	0.00	100.00	5.89	94.12	97.67
ECG Database										
4	0	90.54	90.17	83.26	88.61					92.10
3	1	81.03	82.67	82.08	81.43	7.22	92.78	7.51	92.49	92.74
2	2	68.62	71.62	68.25	69.74	1.16	98.84	12.10	87.9	93.21
1	3	63.34	67.65	67.28	66.73	2.51	97.49	15.56	84.45	93.43

There are only 3 classes present in both the **Iris** and **Wine** datasets. Thus, only one class is used as the novel class. In both cases, the proposed classifier is able to detect most of the novel classes correctly. Moreover, for the **Iris** data set, when there is no novel class present, the proposed classifier shows slightly lower accuracy rates in comparison to those of some of the baseline methods. For the **Wine** data set, the accuracy rate of the proposed hybrid classifier is also slightly decreased when no novel class is added. For both datasets, the accuracy rates of the hybrid classifier are significantly higher than those of other classifiers in the presence of

novel classes. This is owing to the fact that other baseline classifiers are unware of the arrival of novel classes and misclassify them as known categories.

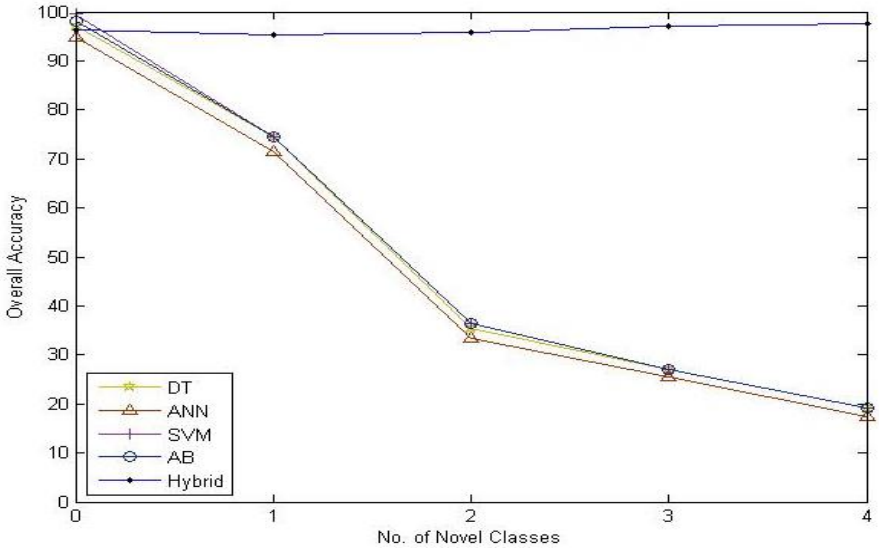


Figure 32 : Performance variations on Glass database
with the increment of the number of novel classes

For the **Glass** dataset, without novel class present, the hybrid classifier achieves slightly lower accuracy rates in comparison to those of other classifiers. However, the overall accuracy increases with the increased number of novel classes as shown in Table 18, whereas the performances of the other classifiers decrease significantly with the increment of the number of novel classes. The accuracy variation along with the increase in the number of novel classes is shown in Figure 32.

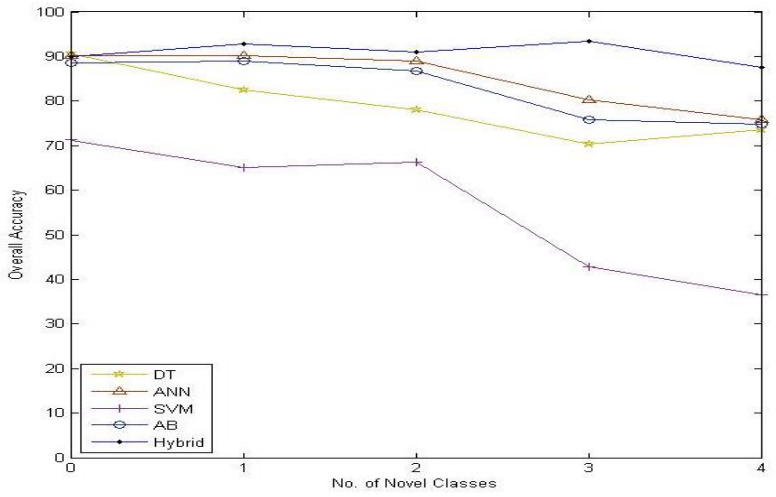


Figure 33 : Performance variations on Soybean Large dataset
with the increment of the number of novel classes

Figure 33 illustrates the performance of **Soybean Large** dataset. As the number of classes present in this dataset is comparatively higher than those of other datasets, increasing the number of novel classes has not changed the accuracy rates as significantly as in other cases since only four novel classes are applied for novel class testing which only represents around 25% of the overall class types in the database. Nevertheless, the overall accuracy of the hybrid classifier is still higher than those of the traditional classifiers in the presence of the novel classes.

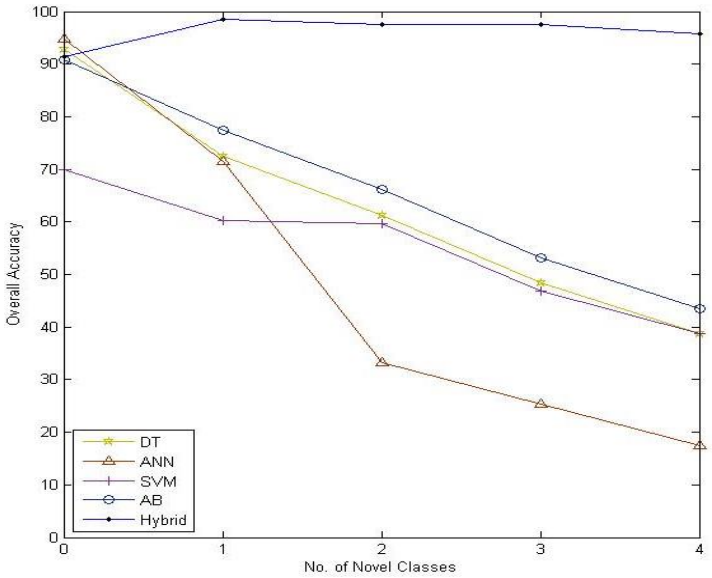


Figure 34 : Performance variations on Image Segmentation dataset with the increment of the number of novel classes

Similarly, for the evaluation of the **Image Segmentation** dataset, the hybrid classifier obtains slightly lower performance in terms of accuracy compared to the results of some baseline classifiers, when there is no novel class present. The performance improves and remains stable with the increase of novel class cases while the performance of other classifiers decreases rapidly as plotted in Figure 34. Interestingly the true positive rate of the hybrid classifier on this dataset is 100% for the testing cases which is also comparatively higher than the results of other methods.

Since the **KDD Cup dataset** is large in size, we have taken a subset of 20,000 random samples from 10 most classes for our experiment. The remaining classes were discarded because of the low instance numbers compared to those from the major classes employed in this research. The performance variations of the hybrid classifier along with other classifiers are illustrated in Figure 35. Similar to other test datasets, the hybrid classifier maintains consistently higher

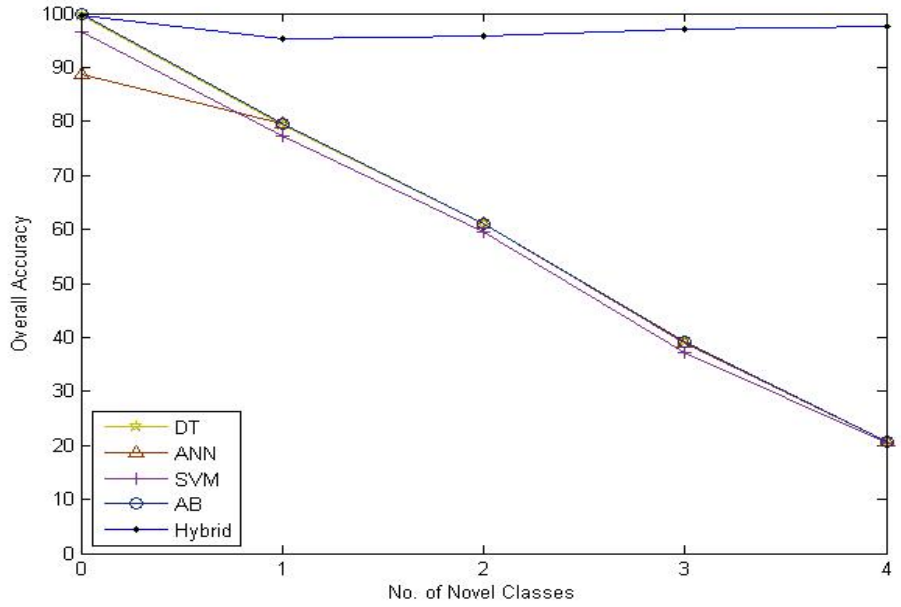


Figure 35 : Performance variations on KDD Cup dataset
with the increment of the number of novel classes

performance with the increase of novel class cases while the accuracies of the other classifiers linearly decrease with the increment of novel class cases.

Figure 36 illustrates the results for the derived **ECG Dataset**. While increasing the number of novel classes, the hybrid classifier is capable of identifying the newly arrived novel classes. However, other baseline classifiers miss-categorize them as known class categories, resulting in the decrement of the overall performance.

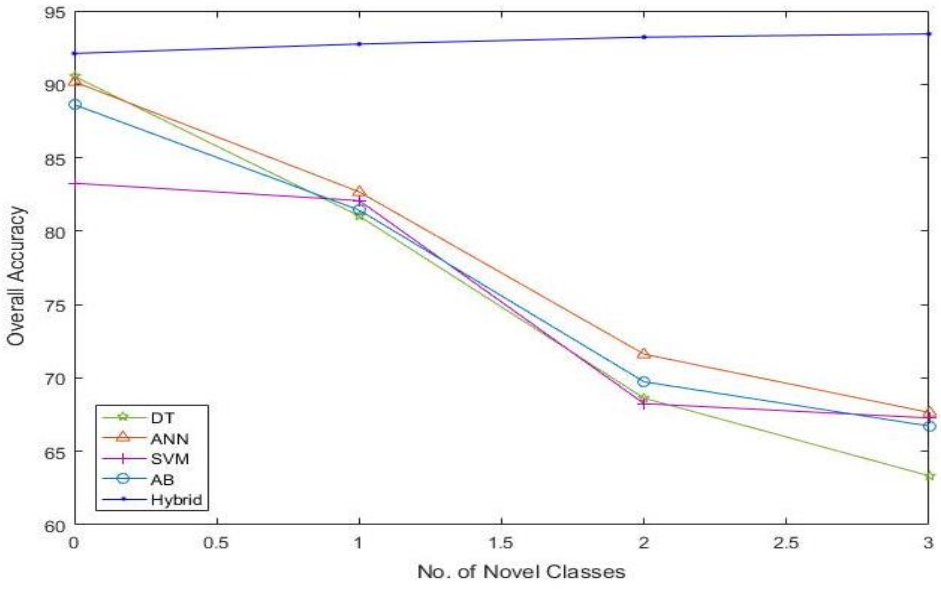


Figure 36 : Performance variations on ECG database
with the increment of the number of novel classes

7.3 Conclusion

In this research, a hybrid ensemble model is proposed for classification along with a novel class detector using combination of multiple techniques. The end result is dependent on the training process and the optimized parameters. A meta-classifier approach is used which gives the freedom of choosing a variety of traditional classifiers to be used as base learners unlike other works in this domain. Our algorithm has been evaluated against some of the widely-used UCI benchmark datasets to prove its performance. The experimental results show its efficiency compared to a few widely used traditional meta classifiers. It also reveals that the classifier can correctly identify novel classes which in turn improves the overall classification accuracy in presence of previously undefined or novel class instances. Although decision trees and neural networks are used as the base learners in the experiments, it can be replaced by any other type of base classifiers. This meta-approach only defines the rules of classifier training and decision-making methods for detecting novel classes. So, the final accuracy might vary depending on the accuracy of chosen base learners.

In future work, the base model of the proposed meta-classifier could be implemented using other classification algorithms such as Artificial Neural Network, SVM etc. to identify its performance variations and classification efficiency. The proposed approach does not consider concept drift property and neither does it attempt to check correlation among multiple instances of the novel class. The proposed method also does not consider multiple types of novel classes but purely combines all of them as a single newly arrived class. So, in future directions, concept evolution or novel class detection could be combined along with concept drifting by adding another layer on top of novel class detection to differentiate multiple novel class types while increasing/maintaining accuracy and building incremental classifier with the arrival of multiple types of novel classes.

Classification is a crucial part of ECG analysis recognizing various abnormality in ECG signals. Given the same sate of features, it is up to the classifier to determine the abnormalities correctly. So, a highly accurate classifier can increase the overall accuracy detecting abnormality. This research proposes a hybrid classifier which could fit into this role aiding heartbeat abnormality detection. Additionally, the proposed novel class detector could be used detecting any type of abnormality that might not be present in the training dataset. A practical use would be labelling those as unknown heartbeats which could later be manually verified rather than incorrectly identified.

Chapter 8 Optimization

The optimization process is not directly related to ECG processing. However, some of the algorithms presented so far (especially meta classifier) require an optimization process (e.g. for optimal parameter selection) to achieve enhanced performance. In this chapter, the optimization algorithm used in the previous chapter is described along with the methodologies and comparison experiments. The firefly algorithm is chosen for the identified optimization tasks, because, as a metaheuristic search method, the algorithm establishes superiority over other swarm intelligence based algorithms for solving diverse optimization problems.

In order to further enhance FA's performance and deal with different optimization problems efficiently and fast, in this research, two improved FA variant algorithms, i.e. Repulsive Firefly Algorithm (RFA) and Scattering Repulsive Firefly Algorithm (SRFA) are proposed which mitigate premature convergence problem of the FA. Besides the original attractiveness movement behaviour of the FA, RFA utilizes a repulsive force strategy to push fireflies away from less optimal regions to reach global optimality in fewer iterations whereas on top of the repulsive force strategy, SRFA also employs a scattering mechanism to divert weak neighbouring solutions to rarely explored distinctive search space to avoid local optimum trap.

Ten benchmark optimization functions are used for the evaluation of the proposed algorithms. The empirical results indicate that RFA and SRFA outperform the original FA and other conventional and state-of-the-art optimization algorithms such as, Simulated Annealing (SA), Particle Swarm Optimization (PSO), Bat Swarm Optimization (BSO), Cuckoo Search Algorithm (CSO), Dragonfly and Ant-Lion Optimization (DFO and ALO), significantly, for the evaluation of all the benchmark functions. A brief review of all the compared algorithm can be found in the work by Yang (2014). Rest of this chapter is divided into three sections. The first section describes the methodologies involved in the development of the algorithm. The next section presents the experiments and finally, the chapter is concluded with remarks and possible future improvement strategies.

8.1 Methodology

In this research, we propose two FA variants, i.e. RFA and SRFA, to mitigate premature convergence problem of the original FA. The former employs the repulsive force strategy while the latter utilizes both the repulsive force and the scattering mechanisms to increase local and global search capabilities.

The repulsive force strategy complements the attractiveness behaviour of the original FA, which accelerates brighter fireflies to move away from less promising search regions drastically to achieve fast convergence. The attractiveness and repulsive search behaviours work alternatively to guide the search in RFA. I.e. when there are brighter fireflies in the neighbourhood, the less bright ones are attracted to the brighter ones by conducting the attractiveness movement as in the original FA. On the contrary, when there are no brighter fireflies in the neighbourhood, instead of conducting purely randomised movement, the current firefly employs the repulsive force action to move away from those with lower light intensities. This repulsive force movement is defined in the following equation.

$$x_i^{t+1} = x_i^t - \beta_0 e^{-\gamma r_{ij}^2}\left(x_j^t - x_i^t\right) \times r_f - \alpha_t \varepsilon_t \tag{8.1}$$

where r_f denotes a repulsive force factor with the assumption that firefly *i* has better fitness than that of firefly *j*. Equation (*8.1*) indicates that the brighter firefly *i* moves away from the less optimal solution *j*. Or the less bright firefly *j* pushes the brighter firefly *i* away from itself. In comparison to the attractiveness behaviour shown in Equation (*3.1*), the repulsive movement enables the current brighter fireflies to move to the opposite direction to those with less brightness to enable fast convergence. Moreover, the repulsive force parameter, r_f, determines the impact of the repulsive force for the position updating of the current firefly *i*. In another word, it determines how much fraction of the repulsive force will be applied to fine tune the repulsive movement.

Figure 37 shows the flowchart of the proposed RFA. It contains the following key steps. First of all, an initial population of fireflies is generated by randomly distributing the fireflies in the search space. Then each firefly is evaluated using the objective function with respect to the problem domain. Next, the fireflies are ranked based on their fitness values and the number of neighbours of each firefly is calculated based on a neighbour distance threshold (dt). Subsequently, the attractiveness and repulsive force movements are used alternatively to move fireflies towards optimal search regions, i.e. either the attractiveness or the repulsive movement is applied based on whether the neighbouring firefly *j* has more or less brightness than the current firefly *i*. I.e. if firefly *j* has better fitness than firefly *i*, firefly *i* moves towards firefly *j* based on the attractive movement as defined in the original FA. Otherwise, if firefly *j* has less fitness than firefly *i*, firefly *i* is pushed away from *j* by employing the repulsive force movement defined in Equation *(8.1)*. A repulsive immune threshold, r_t, is also employed to ensure a certain portion (i.e. 10% of the population) of top ranked fireflies are kept unaffected by the

repulsive force movement; i.e. the immune threshold makes sure a small number of top ranking fireflies are not pushed away from their optimal positions to balance between convergence speed and search diversity. In Figure 37 *n* represents total number of fireflies.

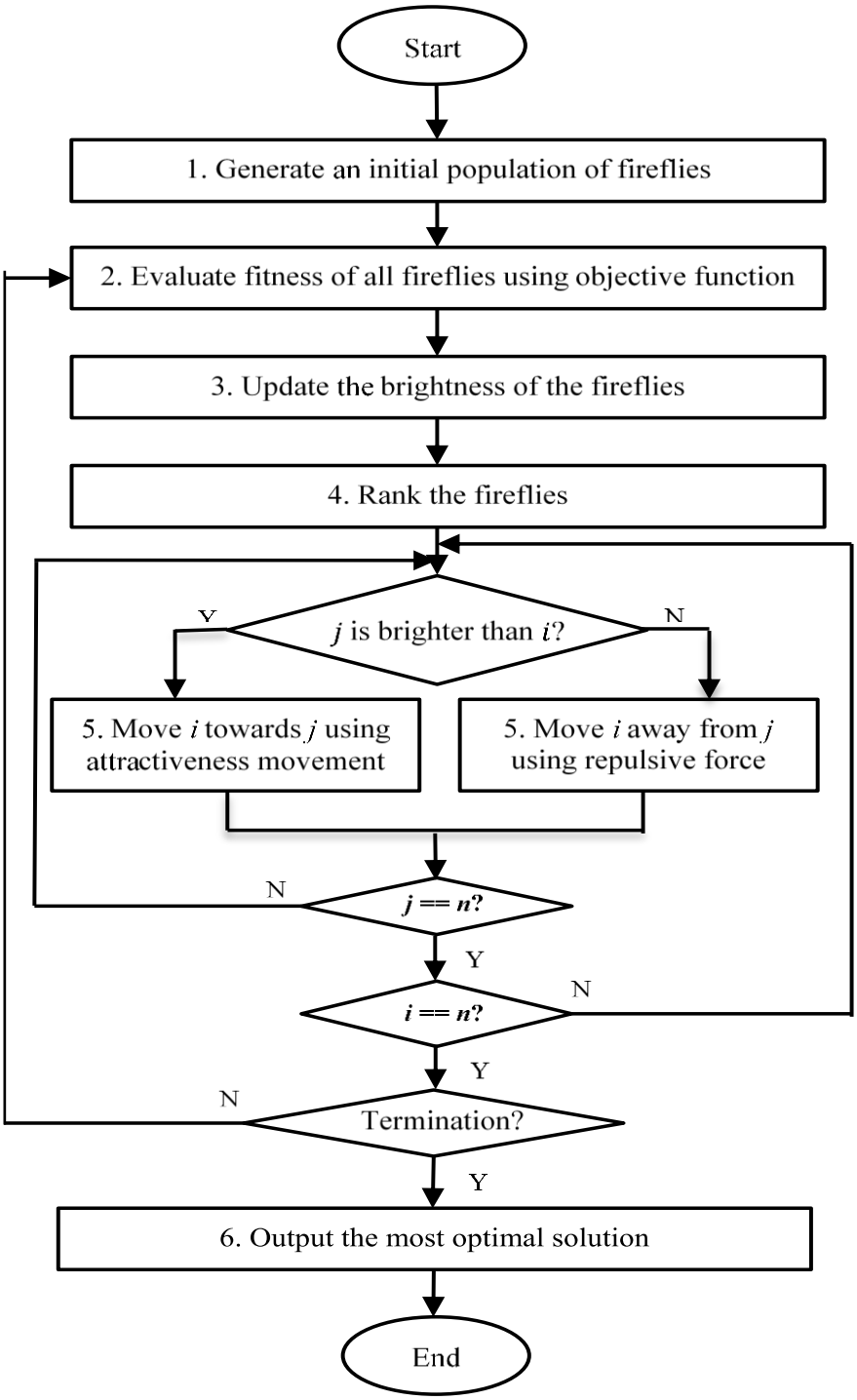


Figure 37 : Flowchart of the proposed RFA

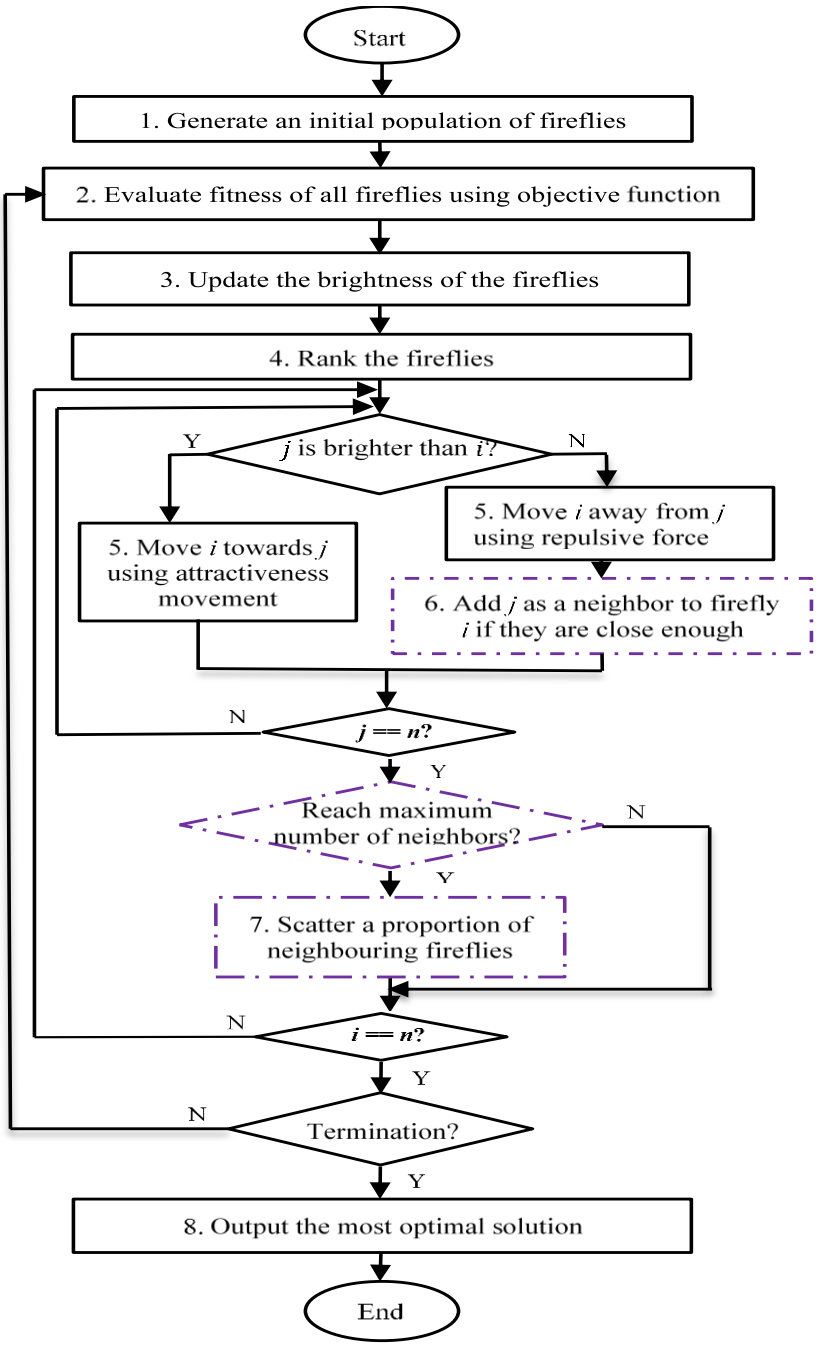


Figure 38 : Flowchart of the proposed SRFA

Overall, in comparison to the original FA where the brighter firefly only conducts random movement, RFA enables the brighter firefly to not only conduct random walk behaviour but also implement the repulsive force action to avoid unpromising search regions and achieve fast convergence.

Another FA variant, SRFA, is also proposed in this research, which incorporates both the repulsive force movement and a scattering mechanism to overcome premature convergence of

the original FA. The scattering mechanism diverts a proportion of weak neighbouring solutions randomly to other distinctive search spaces to increase search diversity. The flowchart of the proposed SRFA is also provided in Figure 38 where the extra steps dedicated to the scattering movement are marked with dotted squares. The pseudocode of the proposed SRFA including both the repulsive force and the scattering actions is provided in Algorithm 9.

As shown in Algorithm 9, a neighbouring distance threshold dt is used in SRFA which states the maximum distance between two fireflies to be considered as neighbours. After employing the repulsive force action to move a brighter firefly *i* away from the less optimal solution *j*, the neighbouring distance threshold is used to determine if the weak firefly *j* is a neighbour of firefly *i*. If it is, the neighbour counting for firefly *i* is increased by 1. In this way, a number of weak solutions are identified as the close neighbours of the current promising solution *i*. Moreover, in order to balance between search diversity and convergence speed, some neighbours (i.e. 5% of the population) of firefly *i* are preserved while others are scatted away to new search regions. A neighbour number threshold nt is therefore used to determine the maximum number of neighbours to keep for firefly i. If the neighbour number exceeds nt, SRFA scatters any extra neighbours randomly to other rarely explored new search regions to increase exploration. In another word, the scattering mechanism keeps a maximum nt number of neighbours for a promising firefly *i* and scatters any extra neighbouring fireflies away to increase search diversity. This random scatter behaviour is defined in the following equation.

$$x_i^d = L_b^d + \mu^d \times (U_b^d - L_b^d) \qquad (8.2)$$

where μ^d is a randomly generated number in the range of [0, 1] for the *d*-th dimension. L_b^d and U_b^d represent the lower and upper bounds in the *d*-th dimension respectively. x_i^d denotes the position of firefly *i* in the *d*-th dimension.

Moreover, the repulsive force action and the scattering behaviour embedded in SRFA work cooperatively to overcome local optima trap. For instance, if the repulsive force action is unable to find a fitter offspring, the scattering search behaviour is able to discover other distinctive regions to widen the search and avoid local optima. On the other hand, if the random scattering is not able to identify more optimal search regions, the repulsive force action enables the brighter fireflies to conduct long jumps and move towards more promising search regions to overcome premature convergence and achieve global optimality.

Algorithm 9: The Scattering Repulsive Firefly Algorithm (SRFA)

Initialization

Objective function: $f(X)$, $X= (x_1, x_2, x_3, x_4, \ldots x_d)$;
Generate an initial population of fireflies X_i (i=1, 2, 3, 4, …, n) ;
Formulate light intensity I so that it is associated with $f(X)$;
Define absorption coefficient γ;
Initialize thresholds as follows;
Default Thresholds

1) Repulsive Force Factor r_f=0.5; //1/2 (default)
2) Repulsive Immune Threshold r_t= n/10; //10% of the population (default)
3) Neighbour Distance Threshold
$$dt = \frac{\sqrt{(U_b - L_b)^2 \times D}}{10}$$
//10% of root square differences in each dimension, where U_b and L_b represent upper and lower bounds for each dimension respectively with D denoting the problem dimensions.
4) Neighbour Number Threshold nt = n/20; //5% of the population (default)

1 **Begin**
2 t = 0;
3 **While (**t < Max_Generation) {
4 S = Zeros (n) ; //S is the array containing scattering flags for each firefly, which are initialized as '0'. S_i represents the scattering flag for firefly i.
5 Vary attractiveness with distance r via exp (-γr^2) ;
6 Evaluate light intensities of the population;
7 Rank the fireflies;
8 **For** i = 1 to n (all n fireflies) {
9 If (S_i = =1) { //Skip if the firefly is marked for scattering
10 Continue; //Go back to line 8 and increase i by 1.
11 }**End If**
12 b_i =0; //Number of neighbours initialized as 0 for firefly i
13 **For** j = 1 to n (all n fireflies) {
14 Calculate distance (r) between firefly i and firefly j;
15 If ($I_j > I_i$) { //Firefly j is brighter than firefly i

16 Move firefly i towards j using Equation (3.1) ;
17 }**Else If (**$i > r_t$**)** { //Repulsive immune threshold
18 Move firefly i away from firefly j using Equation (8.1) ;
19 If ($r < dt$) { //Less than the neighbour distance threshold, dt
20 $b_i = b_i + 1$; //Increase the neighbour counting
21 add j to *NA*; //*NA* is the array containing the indices of neighbouring fireflies
22 }**End If**
23 }**End If**
24 }**End For** j
25 **If (**$b_i > nt$) { //Start of the scattering part (more than the neighbour number threshold, nt)
26 **For** $k = nt$ to b_i {
27 $S_{NA[k]} = 1$;
28 Scatter firefly *NA*[k] using Equation (3) ; //Scatter the k^{th} firefly in *NA*
29 }**End For** k
30 }**End If** //End of the scattering part
31 Clear *NA*;
32 }End For i
33 $t = t + 1$;
34 Rank all the fireflies and find the current global best;
35 }End While
36 Output the most optimal solution;
37 **End**

8.2 Evaluation

A comprehensive evaluation has been conducted to evaluate the proposed two FA variants, SRFA (i.e. RFA with scattering) and RFA (i.e. the repulsive force movement without scattering). In order to indicate the efficiency of the proposed algorithms, several metaheuristic search methods are implemented for performance comparison including PSO, SA, FA, BSO, CSO, DFO and ALO. First experimental setup is discussed for the conducted evaluation.

8.2.1 Experimental Setup

Ten standard benchmark optimization functions were used for evaluation of the proposed algorithms' performance in comparison to other conventional and state-of-the-art search methods. The test functions employed are listed in Table 19, which include various physical properties and shapes. For example, the artificial landscapes 1, 3, 4 and 5 contain many local minima, whereas the artificial landscapes 6, 8, and 9 are bowl-shaped functions and the landscapes 2 and 7 are valley-shaped functions with the test function 10 as a plate-shaped function. These artificial landscapes also represent different optimization problems including multimodal (e.g. 1, 3, 4, and 5) and unimodal optimization problems (e.g. 2, 6, 7, 8, 9, and 10).

Table 19 : Test functions employed in experiments

No.	Name	Formula	Range
1.	Ackley	$f(\mathbf{x}) = -a\exp\left(-b\sqrt{\frac{1}{d}\sum_{i=1}^{d}x_i^2}\right) - \exp\left(\frac{1}{d}\sum_{i=1}^{d}\cos(cx_i)\right) + a + \exp(1)$ $a = 20$, $b = 0.2$ and $c = 2\pi$	[-15, 30]
2.	Dixon-Price	$f(\mathbf{x}) = (x_1 - 1)^2 + \sum_{i=2}^{d} i\left(2x_i^2 - x_{i-1}\right)^2$	[-10, 10]
3.	Griewank	$f(\mathbf{x}) = \sum_{i=1}^{d}\frac{x_i^2}{4000} - \prod_{i=1}^{d}\cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	[-600, 600]
4.	Levy	$f(\mathbf{x}) = \sin^2(\pi w_1) + \sum_{i=1}^{d-1}(w_i - 1)^2\left[1 + 10\sin^2(\pi w_i + 1)\right] + (w_d - 1)^2\left[1 + \sin^2(2\pi w_d)\right]$, where $w_i = 1 + \frac{x_i - 1}{4}$, for all $i = 1, \ldots, d$	[-15, 30]
5.	Rastrigin	$f(\mathbf{x}) = 10d + \sum_{i=1}^{d}\left[x_i^2 - 10\cos(2\pi x_i)\right]$	[-5.12, 5.12]
6.	Rotated Hyper-Ellipsoid	$f(\mathbf{x}) = \sum_{i=1}^{d}\sum_{j=1}^{i}x_j^2$	[-65, 65]
7.	Rosenbrock	$f(\mathbf{x}) = \sum_{i=1}^{d-1}\left[100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2\right]$	[-5,10]
8.	Sphere	$f(\mathbf{x}) = \sum_{i=1}^{d}x_i^2$	[-5.12, 5.12]
9.	Sum of Different Powers	$f(\mathbf{x}) = \sum_{i=1}^{d}\lvert x_i\rvert^{i+1}$	[-1, 1]
10.	Zakharov	$f(\mathbf{x}) = \sum_{i=1}^{d}x_i^2 + \left(\sum_{i=1}^{d}0.5ix_i\right)^2 + \left(\sum_{i=1}^{d}0.5ix_i\right)^4$	[-5, 10]

8.2.2 Evaluation Results

Several experiments are conducted for comparison of the proposed FA variants with other metaheuristic search methods. The first set of experiments conducted focuses on the comparison of convergence speed between different algorithms on selected benchmark functions. Secondly, performance of the algorithms is compared by increasing dimensions exponentially while maintaining a fixed size of population and the number of iterations. Finally, algorithm performance is evaluated for all the test functions under the same (high dimensional) experimental settings.

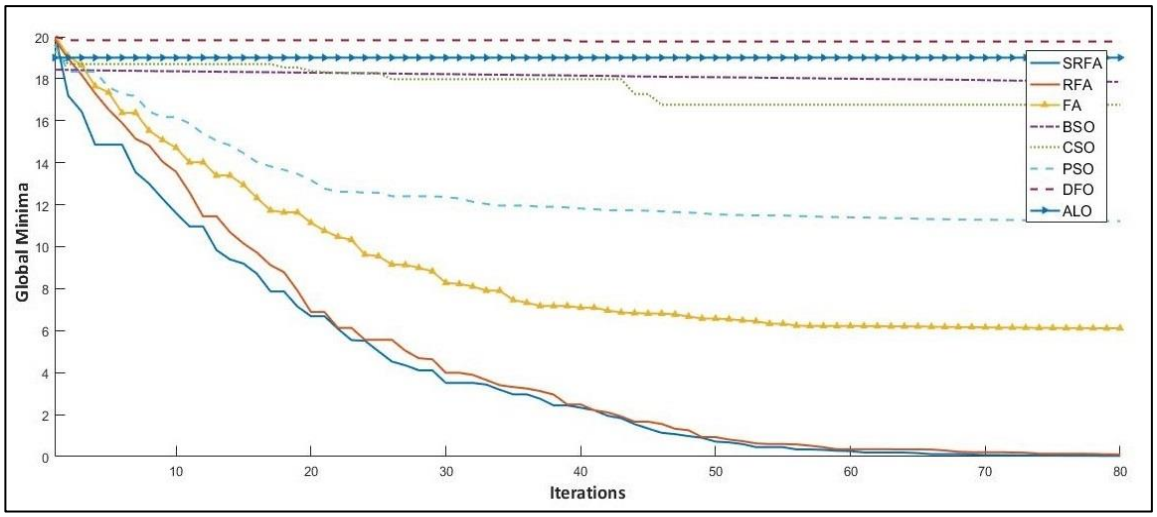


Figure 39 : Convergence curves for the Ackley function

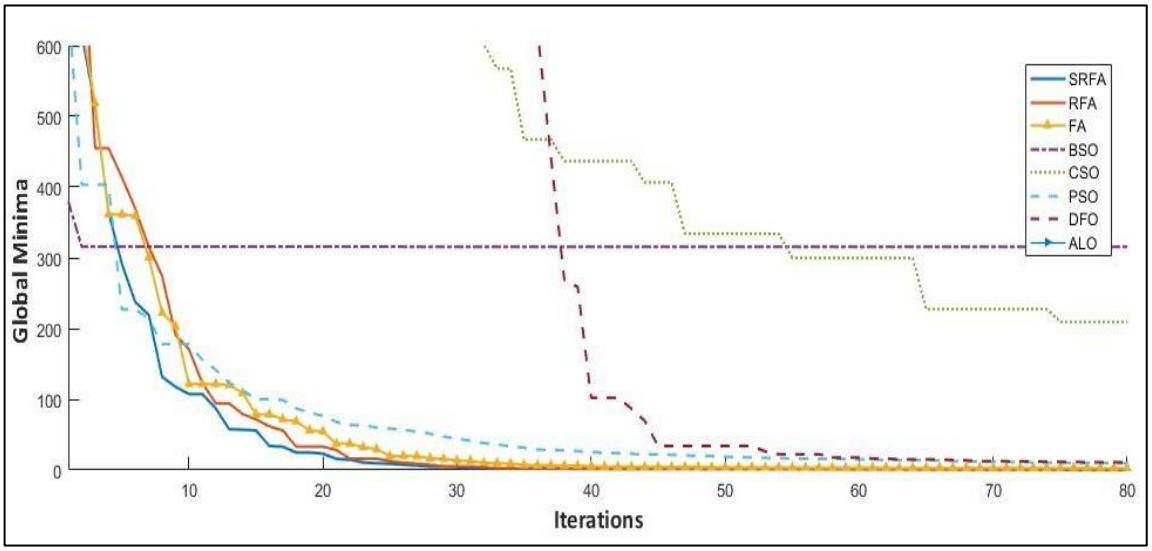


Figure 40 : Convergence curves for the Griewank function

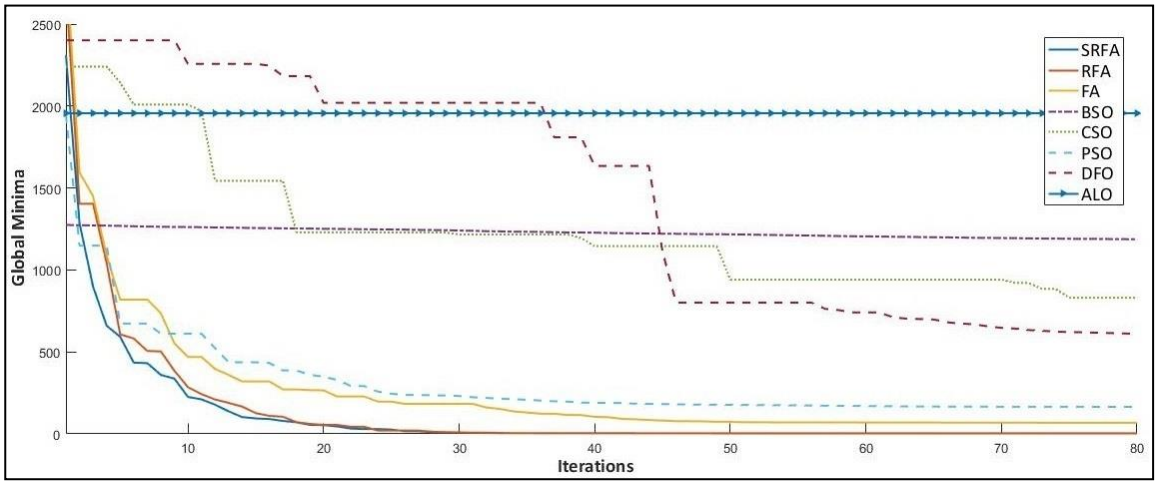


Figure 41 : Convergence curves for the Levy function

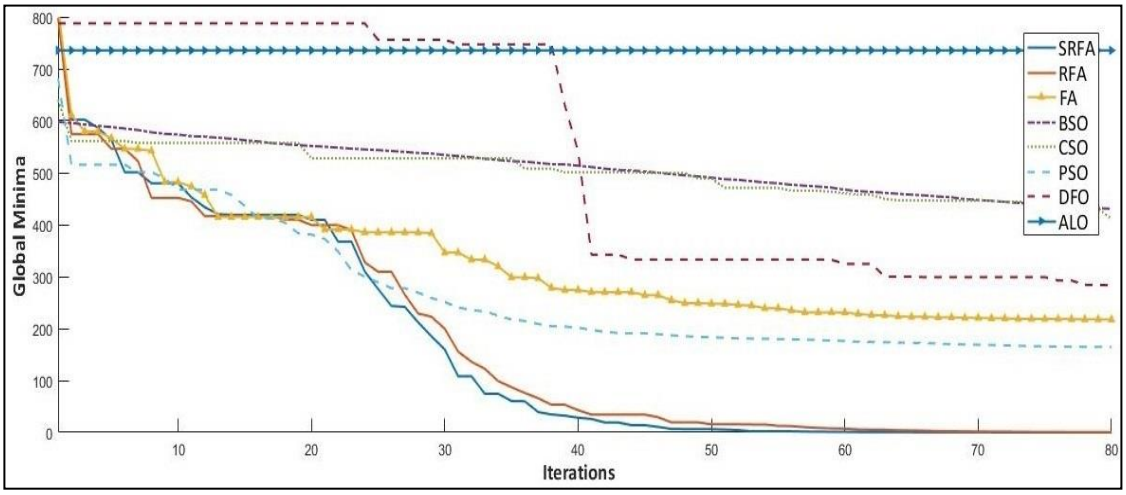


Figure 42 : Convergence curves for the Rastrigin function

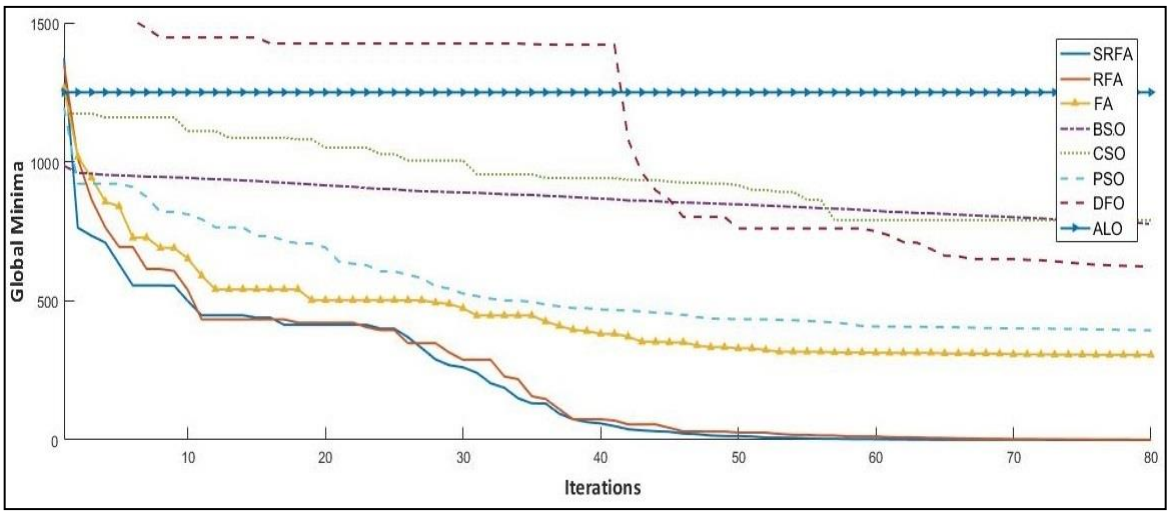


Figure 43 : Convergence curves for the Zakharov function

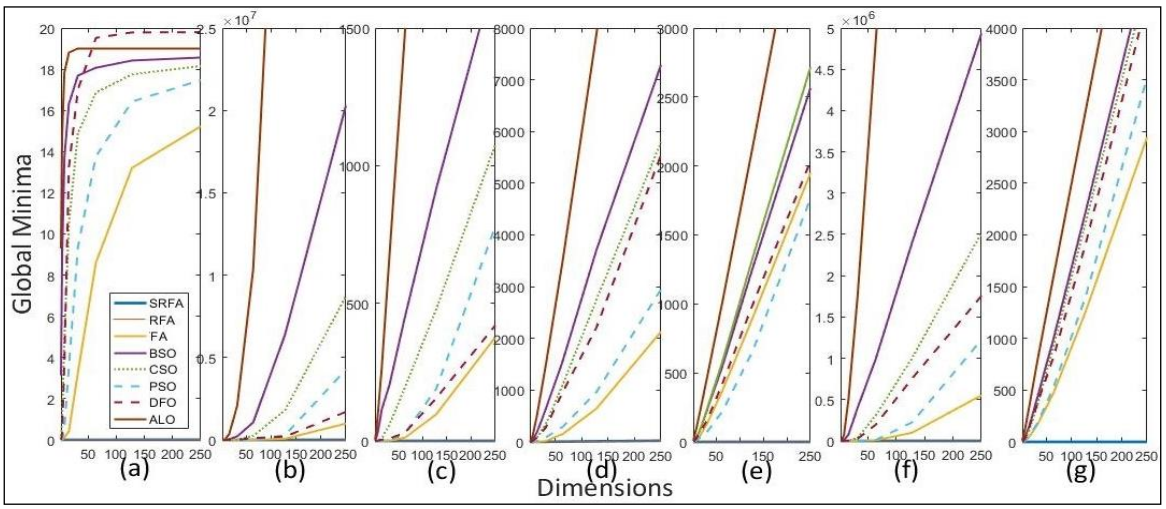


Figure 44 : Final global minima results for all algorithms with the increment of dimensions; (a) Ackley's function, (b) Dixon-Price Function, (c) Griewank Function, (d) Levy function, (e) Rosenbrock function (f) Rastrigin function and (g) Zakharov function.

8.2.2.1 Convergence Comparison

In this experiment, the convergence rates of RFA and SRFA are compared using the ten benchmark functions. This experiment is conducted by keeping the population size and the dimension fixed, i.e. the population size=40 and dimension=50. The empirical results for the evaluation of five functions (i.e. Ackley, Griewank, Levy, Rastrigin and Zakharov) are shown in Figure 39, Figure 40, Figure 41, Figure 42, Figure 43, respectively. These results are obtained from the first 80 iterations. The results of SA are not included since it takes a much larger number of iterations to converge and tends to go beyond the range of the plots for most of the cases. Overall, as illustsrated in Figure 39-Figure 43, RFA and SRFA achieve the fastest convergence for the evaluation of the above five artificial landscapes in most of the cases. A similar observation is also obtained for the evaluation of other optimization functions.

As indicated in Figure 39-Figure 43, both RFA and SRFA show significant convergence improvements in comparison with other methods owing to the proposed repulsive force and scattering stategies. Especially, the repulsive force strategy enables brighter fireflies in the neighbourhood to reach optimal regions in less iterations in comparison to the original FA. Furthermore, SRFA shows a slight convergence improvement in most of the cases in comparison with RFA owing to the scattering mechanism, which diverts a small portion of population congested in a converging location to a random search space to overcome

stagnation. Overall, the results of SRFA are closer to global optima than the outputs of any other algorithms for most of the cases at any given iteration for each test function.

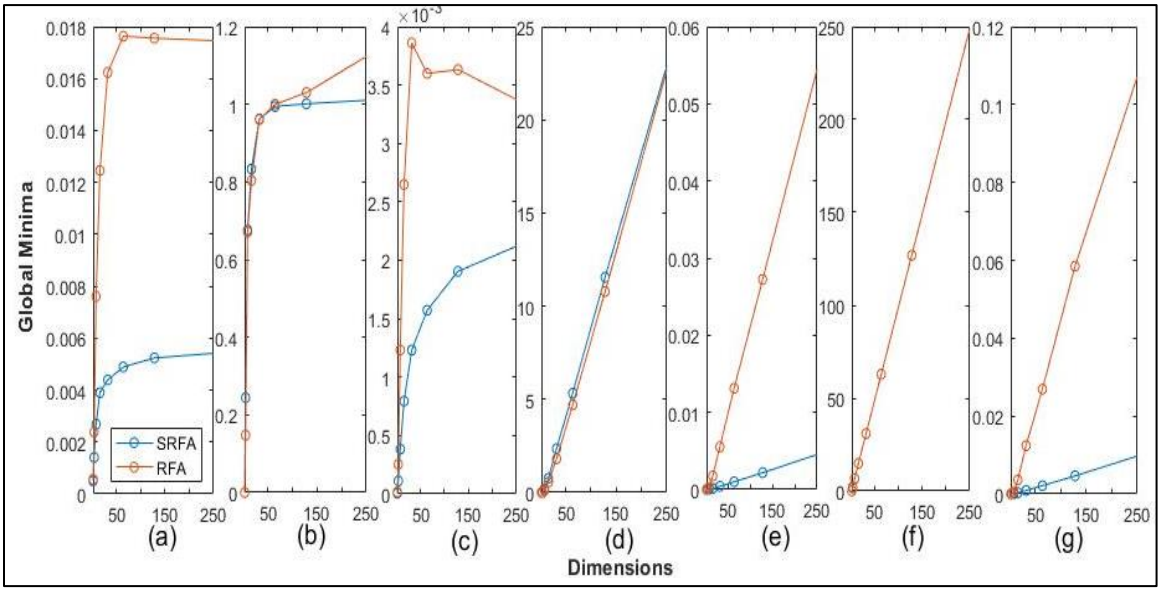


Figure 45 : Final global minima plots scaled for RFA and SRFA with the increment of dimensions with *x* axis representing the dimension which increases exponentially, i.e. 2^i (*i*=1 to 8), and *y* axis representing the global minima ((a) Ackley, (b) Dixon-Price, (c) Griewank, (d) Levy, (e) Rosenbrock, (f) Rastrigin, and (g) Zakharov). The plots of SRFA and RFA are overlapping for the evaluation of the Rastrigin function in (f).

8.2.2.2 Comparison for Problems with Diverse Dimensions

In the second experiment, performance comparison is conducted between different algorithms by increasing dimensions exponentially, i.e. 2^i where i=1 to 8, while maintaining a fixed size of population (i.e. 30 individuals) and the number of iterations (i.e. 100). Specifically, experiments are conducted using dimensions from 2 to 256. For each dimension, all the methods perform 30 runs. The mean value of the minima over the 30 runs for each algorithm for each dimension is used to generate the plots shown in Figure 44 and Figure 45. As shown in Figure 44, the final outcomes of all the related algorithms tend to drift further away from global minima (i.e. zero) while increasing the problem dimension. As the increase of the problem dimensions, the finding of the global optima becomes more and more challenging to other methods and their performances deteriorate more and more drastically. However, increasing dimensions has little impact on the final outcomes of RFA and SRFA. As shown in Figure 44, their plots appear to be overlapping straight lines, very close to the *x* axis, in comparison to those of other methods. To better present the results of both RFA and SRFA with respect to the problem dimension, the outputs of our algorithms are plotted by scaling up

the *y* axis in Figure 45. Overall, SRFA and RFA outperform related methods significantly when the problem dimension increases exponentially. Owing to the cooperation between the repulsive force and the scattering actions, SRFA shows better capabilities in escaping from local optima trap and has the best global minima.

Table 20 : Detailed performance comparison between algorithms using all the benchmark functions

Functions↓		SRFA	RFA	FA	SA	BSO	CSO	PSO	DFO	ALO
Ackley	Mean	5.26E-03	1.45E-02	1.57E+01	2.01E+01	1.90E+01	1.78E+01	1.75E+01	1.98E+01	1.90E+01
	Min	4.14E-03	1.30E-02	1.48E+01	1.94E+01	1.81E+01	1.71E+01	1.69E+01	1.97E+01	1.90E+01
	Max	5.86E-03	1.60E-02	1.66E+01	2.06E+01	1.98E+01	1.83E+01	1.83E+01	1.99E+01	1.90E+01
	Std	3.14E-04	7.32E-04	4.27E-01	2.79E-01	4.29E-01	3.10E-01	2.81E-01	3.92E-02	1.45E-14
Dixon-Price	Mean	1.01E+00	1.05E+00	1.30E+06	1.08E+08	2.02E+07	3.72E+06	3.21E+06	1.12E+06	1.27E+08
	Min	1.00E+00	1.04E+00	7.46E+05	9.60E+07	6.23E+06	1.51E+06	1.47E+06	2.23E+05	1.11E+08
	Max	1.01E+00	1.07E+00	1.90E+06	1.16E+08	4.74E+07	7.84E+06	4.91E+06	5.48E+06	1.45E+08
	Std	6.15E-04	6.76E-03	3.07E+05	4.39E+06	8.95E+06	1.09E+06	8.45E+05	9.99E+05	8.58E+06
Griewank	Mean	2.07E-03	3.36E-03	4.59E+02	4.81E+03	1.80E+03	6.95E+02	6.90E+02	3.60E+02	5.32E+03
	Min	1.56E-03	2.44E-03	3.48E+02	4.59E+03	1.03E+03	5.37E+02	5.17E+02	1.65E+02	4.92E+03
	Max	2.52E-03	4.33E-03	6.04E+02	4.99E+03	3.57E+03	8.56E+02	8.75E+02	6.32E+02	5.90E+03
	Std	2.18E-04	4.46E-04	6.16E+01	1.16E+02	6.05E+02	8.52E+01	9.07E+01	1.10E+02	1.74E+02
Levy	Mean	1.84E+01	1.82E+01	1.98E+03	1.14E+04	6.54E+03	4.23E+03	2.51E+03	5.04E+03	1.35E+04
	Min	1.83E+01	1.78E+01	1.50E+03	7.40E+03	3.88E+03	3.52E+03	1.89E+03	2.85E+03	1.19E+04
	Max	1.86E+01	1.84E+01	2.55E+03	1.26E+04	1.00E+04	4.97E+03	3.20E+03	8.05E+03	1.49E+04
	Std	5.06E-02	1.63E-01	2.53E+02	9.39E+02	1.53E+03	3.77E+02	2.87E+02	1.38E+03	7.41E+02
Rastrigin	Mean	3.46E-03	2.89E-02	1.62E+03	3.24E+03	2.28E+03	2.09E+03	1.42E+03	1.56E+03	3.41E+03
	Min	2.47E-03	2.07E-02	1.36E+03	2.41E+03	1.96E+03	1.90E+03	1.21E+03	9.53E+02	3.23E+03
	Max	4.18E-03	5.26E-02	1.98E+03	3.34E+03	2.85E+03	2.25E+03	1.66E+03	2.07E+03	3.57E+03
	Std	3.44E-04	5.73E-03	1.42E+02	1.62E+02	1.91E+02	6.99E+01	9.88E+01	2.69E+02	8.12E+01
Rotated Hyper-Ellipsoid	Mean	2.81E-01	1.55E+00	1.95E+06	2.19E+07	7.84E+06	3.15E+06	2.87E+06	1.66E+06	2.46E+07
	Min	2.05E-01	1.24E+00	1.41E+06	1.95E+07	4.27E+06	2.23E+06	2.19E+06	8.31E+05	2.18E+07
	Max	3.40E-01	2.08E+00	2.58E+06	2.29E+07	1.27E+07	4.10E+06	3.82E+06	2.93E+06	2.61E+07
	Std	2.83E-02	1.57E-01	2.92E+05	8.46E+05	2.10E+06	4.90E+05	4.23E+05	4.53E+05	8.55E+05
Rosenbrock	Mean	2.12E+02	2.24E+02	2.10E+12	1.75E+14	3.45E+13	1.00E+10	4.82E+12	1.66E+12	2.09E+14
	Min	2.09E+02	2.17E+02	9.37E+11	1.52E+14	6.91E+12	1.00E+10	2.67E+12	1.85E+11	1.78E+14
	Max	2.14E+02	2.29E+02	3.45E+12	1.92E+14	7.32E+13	1.00E+10	8.87E+12	3.93E+12	2.26E+14
	Std	1.43E+00	3.29E+00	6.23E+11	1.04E+13	1.57E+13	0.00E+00	1.47E+12	9.48E+11	1.04E+13
Sphere	Mean	1.78E-05	1.41E-04	1.41E+02	1.40E+03	4.82E+02	2.04E+02	1.98E+02	1.14E+02	1.54E+03
	Min	1.45E-05	1.08E-04	1.13E+02	1.30E+03	2.07E+02	1.44E+02	1.39E+02	3.71E+01	1.45E+03
	Max	2.03E-05	2.09E-04	1.78E+02	1.47E+03	9.80E+02	2.92E+02	2.78E+02	1.64E+02	1.68E+03
	Std	1.35E-06	2.03E-05	1.58E+01	4.55E+01	1.34E+02	3.72E+01	2.93E+01	3.08E+01	5.24E+01
Sum of Different Powers	Mean	1.29E-12	4.81E-12	3.45E-05	2.71E-05	9.94E-03	1.94E-03	1.30E-04	5.56E-02	2.04E+00
	Min	3.37E-15	2.34E-13	2.13E-06	3.61E-06	1.07E-04	1.18E-04	3.00E-06	1.23E-06	1.35E+00
	Max	5.86E-12	1.55E-11	4.38E-04	6.35E-05	6.65E-02	8.11E-03	5.21E-04	3.62E-01	2.86E+00
	Std	1.34E-12	4.52E-12	7.98E-05	1.90E-05	1.86E-02	1.96E-03	1.26E-04	1.01E-01	3.79E-01
Zakharov	Mean	7.66E-03	6.08E-02	2.49E+03	5.79E+03	4.14E+03	3.42E+03	2.85E+03	3.44E+03	5.00E+03
	Min	6.56E-03	3.98E-02	2.15E+03	5.54E+03	3.64E+03	3.16E+03	2.48E+03	2.89E+03	5.00E+03
	Max	8.94E-03	9.16E-02	2.76E+03	5.99E+03	4.68E+03	3.87E+03	3.12E+03	4.29E+03	5.00E+03
	Std	6.28E-04	9.88E-03	1.43E+02	1.20E+02	3.10E+02	1.62E+02	1.70E+02	3.04E+02	0.00E+00

8.2.2.3 Overall Performance Comparison

Some experiments are also conducted to compare the proposed algorithms with all of the related methods for the evaluation of all the benchmark functions. The following settings are applied to each algorithm, i.e. the population size=20, the maximum iterations=100 and the problem dimension=200. The detailed evaluation results for the evaluation of each benchmark function are provided in Table 20. Since all of the algorithms are stochastic methods, 30 runs are evaluated for each algorithm for the evaluation of each optimization function. The mean, minimum, maximum, and standard deviation values obtained from the 30 benchmark runs of each algorithm are used for the evaluation of each artificial landscape. The empirical results indicate that for the evaluation of all the ten benchmark functions, SRFA produces the closest results to the global minima (except for the Levy function where RFA without scattering achieves the most optimal solution). Both SRFA and RFA outperform other conventional and state-of-the-art search methods, significantly.

Some theoretical comparison between the proposed algorithms and other related methods is conducted below. The proposed RFA has employed the repulsive force strategy to identify optimal search regions more effectively than the conventional FA whereas the proposed SRFA integrates both the repulsive force action and the scattering movement to increase local exploitation and global exploration to mitigate premature convergence of the original FA. For instance, if there is no better solution in the neighbourhood, a random exploration is conducted in the original FA. If this random walk action is unable to generate a fitter offspring, this may lead to local optimum trap, whereas the repulsive force action in RFA pushes the current firefly out of gloomy local search regions to reach better positions to overcome stagnation. Moreover, the scattering mechanism and the repulsive force action work cooperatively to drive the search out of local optima in SRFA. For example, when the repulsive force behaviour fails to produce a better offspring, the scattering mechanism re-distributes a proportion of weak neighbouring solutions of the current firefly to other comparatively less exploited search regions to increase the search diversity and reduce the probability of local optima trap. On the contrary, when the scattering movement is unable to produce better solutions, the repulsive force behaviour drives the fireflies out of the unpromising search regions effectively to avoid premature convergence. Therefore, RFA and SRFA outperform the conventional FA and other metaheuristic search methods, significantly, for diverse optimization problems.

Moreover, among the related methods, FA and DFO show better performance than those of other algorithms such as PSO, CSO, ALO, SA, and BSO for the evaluation of the majority of

the test functions. DFO employs static and dynamic swarming behaviours of dragonflies by modelling their social behaviours in searching for food and avoiding enemies and possesses sufficient search diversity than PSO, SA and CSO for finding global optima. Also FA enables each search agent to follow multiple optimal neighbouring solutions to avoid local optima trap and shows great efficiency in dealing with multimodal optimization problems. Overall, FA performs better than DFO for the multimodal optimization functions such as Ackley and Levy, and the unimodal optimization function such as Zakharov whereas DFO achieves better performance than FA for the unimodal optimization functions such as Dixon-Price, Rotated Hyper-Ellipsoid, and Sphere and the multimodal Griewank function. The proposed RFA and SRFA not only preserve but also extend the multimodal search capabilities of the original FA by using the proposed repulsive force and the scattering mechanisms to increase FA's search diversity. Therefore, these characteristics account for the superior performance of RFA and SRFA in comparison to all other methods.

Further experiments were conducted using more complex functions (shifted and rotated versions of the functions described in Table 19 provided in CEC 2014 testsuite. A detailed function definition and evaluation criteria along with sample plots can be found in the provided document (Liang et al. 2013). Three sets of experiments are conducted (with fixed dimention 10, 30 and 50) comparing the other mentioned algorithm. The detailed results for 10, 30 and 40 dimension experiments are included in Appendix II, Appendix III and Appendix IV. In all the experiments, results of 30 runs are used and corresponding minimum, maximum, mean and standard deviation values are compared with other the results of other algorithms. In most of the cases, proposed algorithm performs better than the mentioned algorithms.

8.3 Conclusion

Two FA variant optimization algorithms, RFA and SRFA, were proposed both of which mitigate premature convergence of the original FA. RFA reaches global optimality with a fast convergence rate by utilizing a repulsive force strategy whereas SRFA employs both the repulsive force and scattering mechanisms to increase local and global search capabilities. Evaluated with ten benchmark functions, RFA and SRFA outperform PSO, SA, BSO, CSO, DFO, ALO, and the original FA, significantly, in terms of accuracy and convergence rate, for diverse experimental settings. The proposed algorithm shows great superiority particularly when solving high dimensional optimization problems in comparison to other methods. Although these algorithms achieve impressive performances, the proposed strategies could be

further improved by incorporating diverse chaotic maps (e.g. Gauss and Logistic chaotic maps) and mutation mechanisms (Alguliev et al. 2012; Abdullah et al. 2013) to enable chaotic mutated scattering and repulsive behaviours to further enhance performance.

The proposed optimization algorithm, i.e. RFA, is not directly related to ECG analysis, however, it can be applied in various situations. For example, it could be used for adjusting the parameters of a classifier to increase the accuracy or to optimizing the parameters of a noise reduction algorithm to increase signal to noise ratio. An optimization algorithm like RFA could also help feature selection and feature optimization process before the classification stage. In the previous chapter RFA is used as a supporting algorithm for parameter optimization of hybrid classifier and novel class detector.

Chapter 9 Conclusion

The aim of the work described in this thesis is to identify the challenges of automated ECG processing using only single lead ECG signals and to develop algorithms to deal with the above problems. In this research, various algorithms were proposed and experimented on over different databases. Some of the proposed methods are suitable for online processing on mobile devices while others are dedicated to offline cloud processing. In the next sections, the principle contributions arising from this research are summarised and potential future work and applications are discussed.

9.1 Summary

A number of research contributions raised in this research work are presented in this thesis. Although the research work focuses on abnormality detection using single channel ECG signals, the contributions are distributed into multiple domains. This is because of the range of various methodologies used in multiple stages of ECG processing.

In **Chapter 2**, the components of ECG signal and automated ECG analysis are investigated. The general steps of ECG processing (i.e. pre-processing, feature extraction, feature enhancement and classification) are discussed in detail along with the problems and commonly used solutions.

Chapter 3 explored the current state-of-the-arts in ECG processing. The relevant research was discussed for each key stage of the ECG processing. Most commonly used methods were described in detail. The chapter concluded with current challenges in every corresponding ECG processing stage, especially the challenges of achieving the best trade-off between performance and computational efficiency.

Chapter 4 presented the first research contribution in ECG pre-processing in this work. Multiple techniques were proposed to reduce noise. The first algorithm used wavelet transform and dynamic thresholding processes to increase signal to noise ratio in raw ECG signals. The proposed algorithm was intended for offline use for saved signals. Moreover, the baseline removal technique along with signal smoothing is proposed for preparing raw ECG signals for online processing. The latter was also developed with the intention for the mobile platforms as it is fast and uses only a sliding window of a signal for processing.

Chapter 5 introduced two novel ECG feature extraction schemes. The first method used an innovative strategy to determine the QRS location in ECG signals. The proposed method used

max-min difference curve generation and dynamic thresholding techniques which are highly light weight and do not use a large amount of resources during runtime. Additionally, an R assisted ECG feature extraction scheme was proposed which locates P, Q, S and T locations using a greedy search window without applying any computationally extensive methods (wavelet transform, flourier transform, etc.). Both the methods are sliding window based which makes it possible to work on a small buffer of ECG signals (of a few seconds) stored in the memory and hence makes it suitable for mobile applications. The results from the QRS detection algorithm are compared with the state-of-the-art R-peak detection methods for proving its efficiency in terms of both detection accuracy and the computation complexity.

Furthermore, a standalone feature extraction (P, Q, R, S, and T wave locations) algorithm was also proposed and discussed in this chapter. Later, in the next chapter, those saved results from the above standalone feature extraction algorithm were used to show the opportunity of accuracy improvement using enhanced features.

In **Chapter 6**, a feature enhancement scheme is proposed to improve the accuracy of abnormality detection from the extracted features (from standalone feature extraction algorithm). Two experiments were conducted to prove its efficiency; the first experiment is one with all standard features (relative locations and amplitudes of P, Q, R, S and T waves) and the second experiment involves additional features which add more medically important information on top of the current feature set. The evaluation results show the improvement of accuracy rates for abnormality detection.

Chapter 7 presents a hybrid classification scheme which is capable of working as a general purpose meta classifier. The hybrid classifier was tested against standard baseline ensemble classifiers showing overall improved performance in terms of accuracy.

Furthermore, the proposed hybrid classifier was modified to work as a novel class detector. The novel class detector not only identifies multiple classes present in the training dataset but also detects unknown classes in the test data. The proposed classifier could be used to detect the type of heartbeats which are not present in the training dataset for further analysis. The algorithms are tested using popular datasets from UCI Machine Learning Repository and the results are compared with those of the other commonly used classifiers to indicate its efficiency for novel class detection.

In **Chapter 8**, two novel optimization algorithms (i.e. RFA and SRFA) are presented which are modified versions of the original FA. The proposed algorithms use both attraction and repulsion force actions to reach global optima with fast convergence. The algorithms are used

in the previous chapters for optimal parameter identification in the hybrid classifier. The proposed algorithms are also tested with state-of-the-art FA variants and classical optimization algorithms. The comparison results show that the proposed FA variants achieve better global optima in comparison to those of other FA variants and the original FA.

All of the contributed algorithms work on certain stages of automated ECG processing and improve the overall abnormality and disease detection process. However, as always there are opportunities for further advancement.

9.2 Future directions

In this section, several potential future directions in automated mobile ECG analysis are suggested along with possible challenges.

Deployment of the entire system to mobile platforms will be explored in the future. Such a deployment would involve acquiring the signal from mobile sensors, a real-time analysis of the heartbeats, result recording and real-time disease prediction. However, there will not be many challenges for the deployment of the proposed algorithms to smartphone development kits, as most of the real-time algorithms discussed in this research use light-weight strategies with great computational efficiency. Moreover, more work can also be done on real-time data processing through the Bluetooth channel and sensor optimization techniques (e.g. for mobile ECG sensors).

All of the experiments done in this research used information from multiple stored databases without requiring any participation of human subjects. Most of the databases were acquired more than a few decades ago and mobile ECG sensing has improved significantly in recent years. So, some of the noise reduction mechanisms are not required at all for the new sensors as they come with in-built noise cancellation. In future, live mobile sensor data will be used for testing, instead of using offline databases for the assessment of the proposed system.

Cloud integration is another future direction for mobile ECG processing systems. This can be easily done through smartphones. These days most of the devices are connected to the internet. So, the initial real-time processing can be done using an app in the mobile device and the signal can be sent to a secured server for further processing along with the results of an initial real-time diagnosis. In critical situations emergency services could be notified using the online services. However, there could be many issues. For example, security is one of the concerns, owing to the confidential nature of medical data and the legal requirements for healthcare providers to respect the confidentiality of their patients.

Offline back-end processing can be enabled after cloud integration. Recorded mobile ECG signals (or part of the signals) can be sent to cloud for further processing. This will enable resource consuming methods to operate on disease detection in the cloud environment. Furthermore, ECG processing can be used as a service or a subscribed feature on the cloud by outsourcing all the computationally expensive and complicated processing on the server side.

The other two outcomes of this research are on novel class detection and optimization which are used in ECG processing, however not directly related to ECG analysis. Both of these findings can be used in other application areas. **Novel class detection could be used** in detecting unknown patterns **in multiple domains**. For example, it could be used for detecting unknown network packets (to identify attacks) or it could also be used for the detection of new patterns in data evolving streams. On the other hand, SRFA and RFA can be used in almost every domain to solve high-dimensional optimization problems.

References

Abdullah, A., Deris, S., Anwar, S., and Arjunan, S.N. V, 2013. An Evolutionary Firefly Algorithm for the Estimation of Nonlinear Biological Model Parameters D. J. Klinke, ed. *PLoS ONE*, 8(3), p.e56310.

Addison, P.S., 2005. Wavelet transforms and the ECG: a review. *Physiol. Meas*, 26(5), pp.155–199.

Afonso, V.X., Wieben, O., Tompkins, W.J., and Nguyen, T.Q., 1997. Filter bank-based ECG beat classification. In *Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 'Magnificent Milestones and Emerging Opportunities in Medical Engineering' (Cat. No.97CH36136)*. IEEE, pp. 97–100.

Aha, D.W., Kibler, D., and Albert, M.K., 1991. Instance-Based Learning Algorithms. *Machine Learning*, 6(1), pp.37–66.

Albrecht, P., 1983. *ST segment characterization for long term automated ECG analysis*. M.S. thesis, MIT Dept. of Electrical Engineering and Computer Science.

Alemdar, H., and Ersoy, C., 2010. Wireless sensor networks for healthcare: A survey. *Computer Networks*, 54(15), pp.2688–2710.

Alguliev, R.M., Aliguliyev, R.M., and Isazade, N.R., 2012. DESAMC+DocSum: Differential evolution with self-adaptive mutation and crossover parameters for multi-document summarization. *Knowledge-Based Systems*, 36, pp.21–38.

Alkhaldi, T., Mihaylova, L., and Gellersen, H., 2013. QRS complex detection using centered Cumulative Sums of Squares. In *Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*. IEEE, pp. 168–171.

An-dong, W., Lan, L., and Qin, W., 2012. AASRI An Adaptive Morphologic Filter Applied to ECG De-noising and Extraction of R Peak at Real-time. In *Elsevier,conference*. AASRI Conference on Computational Intelligence and Bioinformatics. pp. 474–479.

Anon, 2016. St.-Petersburg Institute of Cardiological Technics 12-lead Arrhythmia Database. *Https://Physionet.Org/Pn3/Incartdb/*. Available at: [Accessed April 24, 2016].

Arzeno, N.M., Deng, Z. De, and Poon, C.S., 2008. Analysis of first-derivative based QRS detection algorithms. *IEEE Transactions on Biomedical Engineering*, 55(2), pp.478–484.

Babu, A.V., 2013. ECG Signal Analysis Using Data Clustering and Artificial Neural Networks.

International Journal of Emerging Technology and Advanced Engineering, 2(1), pp.82–90.

Bache, K., and Lichman, M., 2013. UCI Machine Learning Repository. *University of California Irvine School of Information*, 2008(14/8), p.0.

Baig, M.M., Gholamhosseini, H., and Connolly, M.J., 2013. A comprehensive survey of wearable and wireless ECG monitoring systems for older adults. *Medical and Biological Engineering and Computing*, 51(5), pp.485–495.

Barnwell, J.D., Klein, J.L., Stallings, C., Sturm, A., Gillespie, M., Fine, J., and Hyslop, W.B., 2012. Image-guided optimization of the ECG trace in cardiac MRI. *International Journal of Cardiovascular Imaging*, 28(3), pp.587–593.

Behar, J., Oster, J., and Clifford, G.D., 2014. Combining and Benchmarking Methods of Foetal ECG Extraction Without Maternal or Scalp Electrode Data. *Physiological Measurement*, 35(8), pp.1569–1589.

Benitez, D.S., Gaydecki, P.A., Zaidi, A., and Fitzpatrick, A.P., 2000. A new QRS detection algorithm based on the Hilbert transform. In *Computers in Cardiology 2000. Vol.27 (Cat. 00CH37163)*. IEEE, pp. 379–382.

Bianchi, A.M., Mendez, M.O., and Cerutti, S., 2010. Processing of signals recorded through smart devices: Sleep-quality assessment. *IEEE Transactions on Information Technology in Biomedicine*, 14(3), pp.741–747.

Borjesson, P.O., Pahlm, O., Sornmo, L., and Nygards, M.-E., 1982. Adaptive QRS Detection Based on Maximum A Posteriori Estimation. *IEEE Transactions on Biomedical Engineering*, BME-29(5), pp.341–351.

Boron, W.F., and Boulpaep, E.L., 2009. *Medical Physiology*, Elsevier India.

Bsoul, M., Minn, H., and Tamil, L., 2011. Apnea MedAssist: Real-time sleep apnea monitor using single-lead ECG. *IEEE Transactions on Information Technology in Biomedicine*, 15(3), pp.416–427.

Buttussi, F., and Chittaro, L., 2008. MOPET: A context-aware and user-adaptive wearable system for fitness training. *Artificial Intelligence in Medicine*, 42(2), pp.153–163.

Ceylan, R., Özbay, Y., and Karlik, B., 2009. A novel approach for classification of ECG arrhythmias: Type-2 fuzzy clustering neural network. *Expert Systems with Applications*, 36(3 PART 2), pp.6721–6726.

Chan, M., Estève, D., Escriba, C., and Campo, E., 2008. A review of smart homes-Present state

and future challenges. *Computer Methods and Programs in Biomedicine*, 91(1), pp.55–81.

Chang, H.T., Chung, C.G., and Chen, M.W., 2013. An e-caring chair for physiological signal measurement and recording. *Medical Engineering and Physics*, 35(2), pp.277–282.

Chiarugi, F., Sakkalis, V., Emmanouilidou, D., Krontiris, T., Varanini, M., and Tollis, I., 2007. Adaptive threshold QRS detector with best channel selection based on a noise rating system. In *Computers in Cardiology*. pp. 157–160.

Chouakri, S.A., Bereksi-Reguig, F., and Taleb-Ahmed, A., 2011. QRS complex detection based on multi wavelet packet decomposition. *Applied Mathematics and Computation*, 217(23), pp.9508–9525.

Chouhan, V.S., and Mehta, S.S., 2008. Detection of QRS Complexes in 12-lead ECG using Adaptive Quantized Threshold. *International Journal of Computer Science and Network Security*, 8(1), pp.155–163.

Christov, I.I., 2004. Real time electrocardiogram QRS detection using combined adaptive threshold. *Biomedical engineering online*, 3(1), p.28.

Chung-Chih, L., Ping-Yeh, L., Po-Kuan, L., Guan-Yu, H., Wei-Lun, L., and Ren-Guey, L., 2008. A Healthcare Integration System for Disease Assessment and Safety Monitoring of Dementia Patients. *Information Technology in Biomedicine, IEEE Transactions on*, 12(5), pp.579–586.

Coyle, S., King-Tong Lau, Moyna, N., O'Gorman, D., Diamond, D., Di Francesco, F., et al., 2010. BIOTEX Biosensing Textiles for Personalised Healthcare Management. *IEEE Transactions on Information Technology in Biomedicine*, 14(2), pp.364–370.

Davenport, C., Cheng, E.Y.L., Kwok, Y.T.T., Lai, A.H.O., Wakabayashi, T., Hyde, C., and Connock, M., 2006. Assessing the diagnostic test accuracy of natriuretic peptides and ECG in the diagnosis of left ventricular systolic dysfunction: A systematic review and meta-analysis. *British Journal of General Practice*, 56(522), pp.48–56.

Dobbs, S., Schmitt, N., and Ozemek, H., 1984. QRS Detection By Template Matching Using Real-Time Correlation On A Microcomputer. *Journal of Clinical Engineering*, 9(3), pp.197–212.

Ebenezer, D., and Krishnamurthy, V., 1993. Wave digital matched filter for electrocardiogram preprocessing. *Journal of Biomedical Engineering*, 15(2), pp.132–134.

Elgendi, M., 2013. Fast QRS Detection with an Optimized Knowledge-Based Method:

Evaluation on 11 Standard ECG Databases. *PLoS ONE*, 8(9), p.e73557.

Elgendi, M., Eskofier, B., and Abbott, D., 2015. Fast T Wave Detection Calibrated by Clinical Knowledge with Annotation of P and T Waves. *Sensors*, 15(7), pp.17693–17714.

Elgendi, M., Eskofier, B., Dokos, S., and Abbott, D., 2014. Revisiting QRS detection methodologies for portable, wearable, battery-operated, and wireless ECG systems L. A. N. Amaral, ed. *PLoS ONE*, 9(1), p.e84018.

Elgendi, M., Mahalingam, S., Jonkman, M., and Boer, F. De, 2008. A Robust QRS Complex Detection Algorithm Using Dynamic Thresholds. In *International Symposium on Computer Science and its Applications*. pp. 153–158.

Elwell, R., and Polikar, R., 2011. Incremental learning of concept drift in nonstationary environments. *IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council*, 22(10), pp.1517–31.

Farahani, S.M., Abshouri, A.A., Nasiri, B., and Meybodi, M.R., 2011. A Gaussian Firefly Algorithm. *International Journal of Machine Learning and Computing*, 1(5), pp.448–453.

Faria, E.R., Gama, J., and Carvalho, A.C., 2013. Novelty detection algorithm for data streams multi-class problems. *Proceedings of the 28th Annual ACM Symposium on Applied Computing*, pp.795–800.

Farid, D.M., and Rahman, C.M., 2012. Novel class detection in concept-drifting data stream mining employing decision tree. In *2012 7th International Conference on Electrical and Computer Engineering, ICECE 2012*. pp. 630–633.

Farid, D.M., Zhang, L., Hossain, A., Rahman, C.M., Strachan, R., Sexton, G., and Dahal, K., 2013. An adaptive ensemble classifier for mining concept drifting data streams. *Expert Systems with Applications*, 40(15), pp.5895–5906.

Friesen, G.M., Jannett, T.C., Jadallah, M.A., Yates, S.L., Quint, S.R., and Nagle, H.T., 1990. A Comparison of the Noise Sensitivity of Nine QRS Detection Algorithms. *IEEE Transactions on Biomedical Engineering*, 37(1), pp.85–98.

Gandomi, A.H., Yang, X.S., Talatahari, S., and Alavi, A.H., 2013. Firefly algorithm with chaos. *Communications in Nonlinear Science and Numerical Simulation*, 18(1), pp.89–98.

Garg, G., Gupta, S., Singh, V., Gupta, J.R.P., and Mittal, A.P., 2011. Identification of optimal wavelet-based algorithm for removal of power line interferences in ECG signals. In *India International Conference on Power Electronics, IICPE 2010*.

Ghaffari, A., Golbayani, H., and Ghasemi, M., 2008. A new mathematical based QRS detector using continuous wavelet transform. *Computers and Electrical Engineering*, 34(2), pp.81–91.

Goldberger, A.L., Amaral, L.A.N., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.-K., and Stanley, H.E., 2000. PhysioBank, PhysioToolkit, and PhysioNet. *Circulation*, 101(23), pp.E215-20.

Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, S.H., 2000. PhysioBank, PhysioToolkit, and PhysioNet Components of a New Research Resource for Complex Physiologic Signals. *Circulation*, 101(23), pp.215–220.

Gothwal, H., Kedawat, S., and Kumar, R., 2011. Cardiac arrhythmias detection in an ECG beat signal using fast fourier transform and artificial neural network. *Journal of Biomedical Science and Engineering*, 4(4), pp.289–296.

H. Zheng, J. Wu, Zheng, H., and Wu, J., 2008. Real-time QRS detection method. In *HealthCom 2008 - 10th International Conference on e-health Networking, Applications and Services*. pp. 169–170.

He, W., Goodkind, D., and Kowal, P., 2016. An Aging World : 2015 International Population Reports. *Aging*, (March), p.165.

Heilman, K.J., Handelman, M., Lewis, G., and Porges, S.W., 2008. Accuracy of the StressEraser® in the detection of cardiac rhythms. *Applied Psychophysiology Biofeedback*, 33(2), pp.83–89.

Hennig, C., Orglmeister, R., and Group, B.E., 2003. QRS Detection Using Zero Crossing Counts. *Progress in Biomedical Research*, 8(3), pp.138–145.

Hii, P.C., and Chung, W.Y., 2011. A comprehensive ubiquitous healthcare solution on an Android™ mobile device. *Sensors*, 11(7), pp.6799–6815.

Hobbs, F.D.R., Bankhead, C., Mukhtar, T., Stevens, S., Perera-Salazar, R., Holt, T., and Salisbury, C., 2016. Clinical workload in UK primary care: a retrospective analysis of 100 million consultations in England, 2007–14. *The Lancet*, 387(10035), pp.2323–2330.

Holland, J.H., 1975. *Adaptation in Natural and Artificial Systems*,

Hou, W., Yang, B., Wuc, C., and Zhou, Z., 2010. RedTrees: A relational decision tree algorithm in streams. *Expert Systems with Applications*, 37(9), pp.6265–6269.

Hsieh, J.C., and Lo, H.C., 2010. The clinical application of a PACS-dependent 12-lead ECG

and image information system in E-medicine and telemedicine. *Journal of Digital Imaging*, 23(4), pp.501–513.

Hu, Y.H., Tompkins, W.J., Urrusti, J.L., and Afonso, V.X., 1993. Applications of artificial neural networks for ECG signal detection and classification. *Journal of electrocardiology*, 26 Suppl, pp.66–73.

Husselmann, A. V, Hawick, K.A., Hawick, P.K., and Science, C., 2012. Parallel Parametric Optimisation with Firefly Algorithms on Graphical Processing Units.

Jen, K., and Hwang, Y., 2008. ECG feature extraction and classification using cepstrum and neural networks. *Journal of Medical and Biological …*, 28(1), pp.31–37.

Johnson, A.E.W., Behar, J., Andreotti, F., Clifford, G.D., and Oster, J., 2015. Multimodal heart beat detection using signal quality indices. *Physiol. Meas.*, 36(8), pp.1665–77.

Jourand, P., De Clercq, H., Corthout, R., and Puers, R., 2009. Textile Integrated Breathing and ECG Monitoring System. *Procedia Chemistry*, 1(1), pp.722–725.

Kabir, M.A., and Shahnaz, C., 2012. Denoising of ECG signals based on noise reduction algorithms in EMD and wavelet domains. *Biomedical Signal Processing and Control*, 7(5), pp.481–489.

Kadambe, S., Murray, R., and Paye Boudreaux-Bartels, G., 1999. Wavelet transform-based QRS complex detector. *IEEE Transactions on Biomedical Engineering*, 46(7), pp.838–848.

Kaplan, G.A., Cohn, B.A., Cohen, R.D., and Guralnik, J., 1988. The decline in ischemic heart disease mortality: prospective evidence from the Alameda County Study. *American journal of epidemiology*, 127(6), pp.1131–1142.

Katakis, I., Tsoumakas, G., and Vlahavas, I., 2006. Dynamic feature space and incremental feature selection for the classification of textual data streams. *Knowledge Discovery from Data Streams*, (March 2014), pp.107–116.

Kaur, M., Singh, B., and Seema, 2011. Comparisons of Different Approaches for Removal of Baseline Wander from ECG Signal. *Response*, pp.30–36.

Kennedy, J., Eberhart, R.C., and Shi, Y., 2001. Swarm Intelligence. *Swarm Intelligence Introduction and Applications*, p.59.

Khan, A.U.R., Othman, M., Madani, S.A., and Khan, S.U., 2014. A survey of mobile cloud computing application models. *IEEE Communications Surveys and Tutorials*, 16(1), pp.393–413.

Khan, M., Aslam, F., Zaidi, T., and Khan, S.A., 2011. Wavelet based ECG denoising using signal-noise residue method. In *5th International Conference on Bioinformatics and Biomedical Engineering, iCBBE 2011*. IEEE, pp. 1–4.

Kim, M.S., Cho, Y.C., Seo, S.T., Son, C.S., and Kim, Y.N., 2011. A new method of ECG feature detection based on combined wavelet transform for u-health service. *Biomedical Engineering Letters*, 1(2), pp.108–115.

Kumar Jaiswal, G., and Paul, R., 2014. Artificial neural network for ecg classification. *Recent Research in Science and Technology*, 6(1), pp.36–38.

Laguna, P., Mark, R.G., Goldberg, a., and Moody, G.B., 1997. A database for evaluation of algorithms for measurement of QT and\nother waveform intervals in the ECG. In *Computers in Cardiology 1997*. pp. 673–676.

Lee, J., Jeong, K., Yoon, J., and Lee, M., 1996. A simple real-time QRS detection algorithm. In *Proceedings of 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, pp. 1396–1398.

Lee, Y.D., and Chung, W.Y., 2009. Wireless sensor network based wearable smart shirt for ubiquitous health and activity monitoring. *Sensors and Actuators, B: Chemical*, 140(2), pp.390–395.

Li, Z., Ni, J., and Gu, X., 2012. A denoising framework for ECG signal preprocessing. In *Proceedings - 6th International Conference on Internet Computing for Science and Engineering, ICICSE 2012*. IEEE, pp. 176–179.

Liang, J.J., Qu, B.Y., Suganthan, P.N., and Hernández-Díaz, A.G., 2013. Problem Definitions and Evaluation Criteria for the CEC 2013 Special Session on Real-Parameter Optimization.

Lin, C., Mailhes, C., and Tourneret, J.Y., 2010. P- and T-wave delineation in ECG signals using a bayesian approach and a partially collapsed gibbs sampler. *IEEE Transactions on Biomedical Engineering*, 57(12), pp.2840–2849.

Lin, C.F., 2012. Mobile telemedicine: A survey study. *Journal of Medical Systems*, 36(2), pp.511–520.

Lin, C.T., Chang, K.C., Lin, C.L., Chiang, C.C., Lu, S.W., Chang, S.S., Lin, B.S., Liang, H.Y., Chen, R.J., Lee, Y.T., and Ko, L.W., 2010. An intelligent telecardiology system using a wearable and wireless ecg to detect atrial fibrillation. *IEEE Transactions on Information Technology in Biomedicine*, 14(3), pp.726–733.

López, G., Custodio, V., and Moreno, J.I., 2010. LOBIN: E-textile and wireless-sensor-network-based platform for healthcare monitoring in future hospital environments. *IEEE Transactions on Information Technology in Biomedicine*, 14(6), pp.1446–1458.

Madeiro, J.P. V, Cortez, P.C., Marques, J.A.L., Seisdedos, C.R. V, and Sobrinho, C.R.M.R., 2012. An innovative approach of QRS segmentation based on first-derivative, Hilbert and Wavelet Transforms. *Medical Engineering and Physics*, 34(9), pp.1236–1246.

Magjarevic, R., Nagel, J.H., Golzan, S.M., Hakimpour, F., and Toolou, A., 2009. Fetal ECG Extraction Using Multi-Layer Perceptron Neural Networks with Bayesian Approach J. Vander Sloten et al., eds. *4th European Conference of the International Federation for Medical and Biological Engineering*, 22, pp.1378–1385.

Mai, V., Khalil, I., and Meli, C., 2011. ECG biometric using multilayer perceptron and radial basis function neural networks. *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp.2745–2748.

Mamaghanian, H., Khaled, N., Atienza, D., and Vandergheynst, P., 2011. Compressed sensing for real-time energy-efficient ECG compression on wireless body sensor nodes. *IEEE Transactions on Biomedical Engineering*, 58(9), pp.2456–2466.

Markou, M., and Singh, S., 2003a. Novelty detection: A review - Part 1: Statistical approaches. *Signal Processing*, 83(12), pp.2481–2497.

Markou, M., and Singh, S., 2003b. Novelty detection: A review - Part 2:: Neural network based approaches. *Signal Processing*, 83(12), pp.2499–2521.

Martínez, J.P., Almeida, R., Olmos, S., Rocha, A.P., and Laguna, P., 2004. A Wavelet-Based ECG Delineator Evaluation on Standard Databases. *IEEE Transactions on Biomedical Engineering*, 51(4), pp.570–581.

Masud, M., Gao, J., Khan, L., Han, J., and Thuraisingham, B.M., 2011. Classification and novel class detection in concept-drifting data streams under time constraints. *IEEE Transactions on Knowledge and Data Engineering*, 23(6), pp.859–874.

Masud, M.M., Chen, Q., Khan, L., Aggarwal, C.C., Gao, J., Han, J., Srivastava, A., and Oza, N.C., 2013. Classification and adaptive novel class detection of feature-evolving data streams. *IEEE Transactions on Knowledge and Data Engineering*, 25(7), pp.1484–1497.

MathWorks, 2016. Makers of MATLAB and Simulink. Available at: [Accessed March 28, 2017].

Messaoud, M.B., Khelil, B., and Kachouri, A., 2009. Analysis and parameter extraction of P

wave using correlation method. *International Arab Journal of Information Technology (IAJIT)*, 6(1), pp.40–46.

Miao, Y., Qiu, L., Chen, H., Zhang, J., and Wen, Y., 2013. Novel Class Detection within Classification for Data Streams. In *Proceedings of the 10th international conference on Advances in Neural Networks - Volume Part II*. Springer-Verlag, pp. 413–420.

Moavenian, M., and Khorrami, H., 2010. A qualitative comparison of Artificial Neural Networks and Support Vector Machines in ECG arrhythmias classification. *Expert Systems with Applications*, 37(4), pp.3088–3093.

Moody, G.B., and Mark, R.G., 2001. The impact of the MIT-BIH arrhythmia database. *IEEE Engineering in Medicine and Biology Magazine*, 20(3), pp.45–50.

Moraes, J., Freitas, M.M., Vilani, F.N., and Costa, E.V., 2002. A QRS complex detection algorithm using electrocardiogram leads. In *Computers in Cardiology*. pp. 205–208.

Mozaffarian, D., Benjamin, E., Go, A., Arnet, D., Blaha, M., Cushman, M., et al., 2015. Heart Disease and Stroke Statistics – At-a-Glance Heart Disease , Stroke and other Cardiovascular Diseases Heart Disease , Stroke and Cardiovascular Disease Risk Factors. *American Heart Association*, (1), pp.7–10.

Murty, M.N., and Devi, V.S., 2011. Bayes Classifier. In Springer London, pp. 86–102.

Nason, G.P., and Silverman, B.W., 1995. The Stationary Wavelet Transform and some Statistical Applications. *Wavelets and statistics*, pp.281–299.

Oresko, J.J., Jin, Z., Cheng, J., Huang, S., Sun, Y., Duschl, H., and Cheng, A.C., 2010. A wearable smartphone-based platform for real-time cardiovascular disease detection via electrocardiogram processing. *IEEE Transactions on Information Technology in Biomedicine*, 14(3), pp.734–740.

Özbay, Y., and Tezel, G., 2010. A new method for classification of ECG arrhythmias using neural network with adaptive activation function. *Digital Signal Processing*, 20(4), pp.1040–1049.

Pan, J., and Tompkins, W.J., 1985. A Real-Time QRS Detection Algorithm. *Biomedical Engineering, IEEE Transactions on DOI - 10.1109/TBME.1985.325532*, BME-32(3), pp.230–236.

Pan, T., Zhang, L., and Zhou, S., 2010. Detection of ECG characteristic points using biorthogonal spline wavelet. *Proceedings - 2010 3rd International Conference on Biomedical Engineering and Informatics, BMEI 2010*, 2(1), pp.858–863.

Pandian, P.S., Mohanavelu, K., Safeer, K.P., Kotresh, T.M., Shakunthala, D.T., Gopal, P., and Padaki, V.C., 2008. Smart Vest: Wearable multi-parameter remote physiological monitoring system. *Medical Engineering and Physics*, 30(4), pp.466–477.

Pandit, D., Zhang, L., Aslam, N., Liu, C., and Chattopadhyay, S., 2016. Improved abnormality detection from raw ECG signals using feature enhancement. In *2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery, ICNC-FSKD 2016*. IEEE, pp. 1402–1406.

Pandit, D., Zhang, L., Aslam, N., Liu, C., Hossain, A., and Chattopadhyay, S., 2014. An efficient abnormal beat detection scheme from ECG signals using neural network and ensemble classifiers. In *The 8th International Conference on Software, Knowledge, Information Management and Applications (SKIMA 2014)*. IEEE, pp. 1–6.

Pandit, D., Zhang, L., Liu, C., Aslam, N., Chattopadhyay, S., and Lim, C.P., 2017. Noise Reduction in ECG Signals Using Wavelet Transform and Dynamic Thresholding. In Springer Singapore, pp. 193–206.

Pandit, D., Zhang, L., Liu, C., Chattopadhyay, S., Aslam, N., and Lim, C.P., 2017. A lightweight QRS detector for single lead ECG signals using a max-min difference algorithm. *Computer Methods and Programs in Biomedicine*, 144, pp.61–75.

Pinsky, M., 2009. *Introduction to Fourier analysis and wavelets*, American Mathematical Society.

Poli, S., Barbaro, V., Bartolini, P., Calcagnini, G., and Censi, F., 2003. Prediction of atrial fibrillation from surface ECG: Review of methods and algorithms. *Annali dell'Istituto Superiore di Sanita*, 39(2), pp.195–203.

Pollonini, L., Rajan, N.O., Xu, S., Madala, S., and Dacso, C.C., 2012. A novel handheld device for use in remote patient monitoring of heart failure patients-design and preliminary validation on healthy subjects. *Journal of Medical Systems*, 36(2), pp.653–659.

Puurtinen, M., Viik, J., and Hyttinen, J., 2009. Best electrode locations for a small bipolar ECG device: Signal strength analysis of clinical data. *Annals of Biomedical Engineering*, 37(2), pp.331–336.

Rahimpour, M., and Mohammadzadeh Asl, B., 2016. *P* wave detection in ECG signals using an extended Kalman filter: an evaluation in different arrhythmia contexts. *Physiological Measurement*, 37(7), pp.1089–1104.

Rangayyan, R.M., 2002. Biomedical Signal Analysis: A Case-Study Approach. *Signals*, 30(7),

p.552.

Rani, S., Kaur, A., and Ubhi, J.S., 2011. Comparative study of FIR and IIR filters for the removal of Baseline noises from ECG signal. *International Journal of Computer Science and Information Technologies*, 2(3), pp.1105–1108.

Rao, K.D., 2015. DWT Based Detection of R-peaks and Data Compression of ECG Signals. *IETE Journal of Research*, 43(5), pp.345–349.

Raphisak, P., Schuckers, S.C., and Curry, a. D.J., 2004. An algorithm for EMG noise detection in large ECG data. *Computers in Cardiology, 2004*, 1(1), pp.369–372.

RapidMiner, 2016. Predictive Analytics Platform | RapidMiner.

Rezk, S., Asmi, S. El, Poincar, H., and Nancy, I., 2011. an Algebraic Derivative-Based Method for R Wave Detection. In *European Signal Processing Conference*. pp. 1578–1582.

Rodríguez, R., Mexicano, A., Bila, J., Cervantes, S., and Ponce, R., 2015. Feature Extraction of Electrocardiogram Signals by Applying Adaptive Threshold and Principal Component Analysis. *Journal of Applied Research and Technology*, 13(2), pp.261–269.

Rokach, L., 2010. Ensemble-based classifiers. *Artificial Intelligence Review*, 33(1–2), pp.1–39.

Safavian, S.R., and Landgrebe, D., 1991. A Survey of Decision Tree Classifier Methodology. *IEEE Transactions on Systems, Man and Cybernetics*, 21(3), pp.660–674.

Sankari, Z., and Adeli, H., 2011. HeartSaver: A mobile cardiac monitoring system for auto-detection of atrial fibrillation, myocardial infarction, and atrio-ventricular block. *Computers in Biology and Medicine*, 41(4), pp.211–220.

Dos Santos Coelho, L., De Andrade Bernert, D.L., and Mariani, V.C., 2011. A chaotic firefly algorithm applied to reliability-redundancy optimization. In *2011 IEEE Congress of Evolutionary Computation, CEC 2011*. pp. 517–521.

Sardini, E., and Serpelloni, M., 2010. Instrumented wearable belt for wireless health monitoring. *Procedia Engineering*, 5, pp.580–583.

Scherr, D., Dalal, D., Henrikson, C.A., Spragg, D.D., Berger, R.D., Calkins, H., and Cheng, A., 2008. Prospective comparison of the diagnostic utility of a standard event monitor versus a 'leadless' portable ECG monitor in the evaluation of patients with palpitations. *Journal of Interventional Cardiac Electrophysiology*, 22(1), pp.39–44.

Schölkopf, B., 1998. SVMs - A practical consequence of learning theory. *IEEE Intelligent Systems and Their Applications*, 13(4), pp.18–21.

Smith, T.C., and Frank, E., 2016. Introducing machine learning concepts with WEKA. In *Methods in Molecular Biology*. pp. 353–378.

Spinosa, E.J., de Leon F. de Carvalho, A.P., and Gama, J., 2007. OLINDDA: A cluster-based approach for detecting novelty and concept drift in data streams. In *Proceedings of the 2007 ACM symposium on Applied computing - SAC '07*. New York, New York, USA: ACM Press, p. 448.

Spinosa, E.J., de Leon F. de Carvalho, A.P., Gama, J., and De Carvalho, A.P.D.L.F., 2008. Cluster-based novel concept detection in data streams applied to intrusion detection in computer networks. In *Proceedings of the 2008 ACM symposium on Applied computing*. pp. 976–980.

Strintzis, M.G., Stalidis, G., Magnisalis, X., and Maglaveras, N., 1992. Use of Neural Networks for Electrocarciogram (ECG) Feature Extraction, Recognition and Classification. *Neural Network World Journal*, 3(4), pp.313–328.

Subutic, M., Tuba, M., and Stanarevic, N., 2012. Parallelization of the firefly algorithm for unconstrained optimization problems. *Latest Advances in Information …*, pp.264–269.

Taddei, A., Distante, G., Emdin, M., Pisani, P., Moody, G.B., Zeelenberg, C., and Marchesi, C., 1992. The European ST-T database: standard for evaluating systems for the analysis of ST-T changes in ambulatory electrocardiography. *European heart journal*, 13(9), pp.1164–72.

Tadejko, P., and Rakowski, W., 2007. Mathematical morphology based ECG feature extraction for the purpose of heartbeat classification. In *Proceedings - 6th International Conference on Computer Information Systems and Industrial Management Applications, CISIM 2007*. pp. 322–327.

Tamil, E.M., Kamarudin, N.H., Salleh, R., Idris, M.Y.I., Noor, M., and Tamil, A.M., 2008. Heartbeat Electrocardiogram (ECG) signal feature extraction using Discrete Wavelet Transforms (DWT). In *Cspa (2008)*. pp. 1112–1117.

Tay, F.E.H., Guo, D.G., Xu, L., Nyan, M.N., and Yap, K.L., 2009. MEMSWear-biomonitoring system for remote vital signs monitoring. *Journal of the Franklin Institute*, 346(6), pp.531–542.

Tayel, M.B., and El-Bouridy, M.E., 2008. ECG images classification using artificial neural network based on several feature extraction methods. In *2008 International Conference on Computer Engineering & Systems*. pp. 113–115.

Thalkar, S., 2013. Various techniques for removal of power line interference from ECG signal. *International Journal of Scientific & Engineering Research*, 4(12), pp.12–23.

Tilahun, S.L., and Ong, H.C., 2012. Modified firefly algorithm. *Journal of Applied Mathematics*, 2012, pp.1–12.

Tiwari, R., and Dubey, R., 2014. Analysis of dfferent denoising techniques of ECG signals. *International Journal of Emerging Technology and Advanced Engineering*, 4(3), pp.2–6.

Tscherning, H., 2011. Mobile Devices in Social Contexts. *Library Technology Reports*, 47(2), pp.11–23.

Ubeyli, E.D., 2008. Feature extraction for analysis of ECG signals. In *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*. pp. 1080–1083.

Übeyli, E.D., 2008. Implementing wavelet transform/mixture of experts network for analysis of electrocardiogram beats. *Expert Systems*, 25(2), pp.150–162.

Ullah, S., Higgins, H., Braem, B., Latre, B., Blondia, C., Moerman, I., Saleem, S., Rahman, Z., and Kwak, K.S., 2012. A comprehensive survey of wireless body area networks on PHY, MAC, and network layers solutions. *Journal of Medical Systems*, 36(3), pp.1065–1094.

Vijaya, G., Kumar, V., and Verma, H.K., 1998. ANN-based QRS-complex analysis of ECG. *Journal of medical engineering & technology*, 22(4), pp.160–7.

Virgilio, V. Di, Francaiancia, C., Lino, S., and Cerutti, S., 1995. ECG fiducial points detection through wavelet transform. In *Proceedings of 17th International Conference of the Engineering in Medicine and Biology Society*. pp. 747–827.

Wenerstrom, B., and Giraud-Carrier, C., 2006. Temporal data mining in dynamic feature spaces. In *Proceedings - IEEE International Conference on Data Mining, ICDM*. IEEE, pp. 1141–1145.

WHO, 2015. WHO _ Cardiovascular diseases (CVDs). *Cardiovascular diseases (CVDs)*.

Winterhalter, M., Schiller, J., Münte, S., Bund, M., Hoy, L., Weilbach, C., Piepenbrock, S., and Rahe-Meyer, N., 2008. Prospective investigation into the influence of various stressors on skin impedance. *Journal of Clinical Monitoring and Computing*, 22(1), pp.67–74.

Wu, W.H., Bui, A.A.T., Batalin, M.A., Au, L.K., Binney, J.D., and Kaiser, W.J., 2008. MEDIC: Medical embedded device for individualized care. *Artificial Intelligence in*

Medicine, 42(2), pp.137–152.

Xue, Q., Hu, Y.H., and Tompkins, W.J., 1992. Neural-Network-Based Adaptive Matched Filtering for QRS Detection. *IEEE Transactions on Biomedical Engineering*, 39(4), pp.317–329.

Yang, X.-S., 2014. *Nature-Inspired Optimization Algorithms*,

Yang, X.S., 2012. Efficiency analysis of swarm intelligence and randomization techniques. *Journal of Computational and Theoretical Nanoscience*, 9(2), pp.189–198.

Yang, X.S., 2010. Firefly algorithm, Levy flights and global optimization. *Research and Development in Intelligent Systems XXVI: Incorporating Applications and Innovations in Intelligent Systems XVII*, pp.209–218.

Yang, X.S., 2009. Firefly algorithms for multimodal optimization. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. pp. 169–178.

Yang, X.S., and Deb, S., 2009. Cuckoo search via Levy flights. *2009 World Congress on Nature and Biologically Inspired Computing, NABIC 2009 - Proceedings*, pp.210–214.

Yeh, Y.C., and Wang, W.J., 2008. QRS complexes detection for ECG signal: The Difference Operation Method. *Computer Methods and Programs in Biomedicine*, 91(3), pp.245–254.

Ying-Wen Bai, Wen-Yang Chu, Chien-Yu Chen, Yi-Ting Lee, Yi-Ching Tsai, and Cheng-Hung Tsai, 2004. Adjustable 60Hz noise reduction by a notch filter for ECG signals. In *Proceedings of the 21st IEEE Instrumentation and Measurement Technology Conference (IEEE Cat. No.04CH37510)*. pp. 1706–1711.

Yoo, J., Yan, L., Lee, S., Kim, H., Kim, B., and Yoo, H.J., 2009. An attachable ECG sensor bandage with planar-fashionable circuit board. In *Proceedings - International Symposium on Wearable Computers, ISWC*. IEEE, pp. 145–146.

Yoo, J., Yan, L., Lee, S., Kim, H., and Yoo, H.J., 2009. A wearable ECG acquisition system with compact planar-fashionable circuit board-based shirt. *IEEE Transactions on Information Technology in Biomedicine*, 13(6), pp.897–902.

Yoon, Y., Cho, J.H., and Yoon, G., 2009. Non-constrained blood pressure monitoring using ECG and PPG for personal healthcare. *Journal of Medical Systems*, 33(4), pp.261–266.

ZareMoodi, P., Beigy, H., and Kamali Siahroudi, S., 2015. Novel class detection in data streams using local patterns and neighborhood graph. *Neurocomputing*, 158, pp.234–245.

Zhang, G.P., 2000. Neural networks for classification: A survey. *IEEE Transactions on*

Systems, Man and Cybernetics Part C: Applications and Reviews, 30(4), pp.451–462.

Zhang, Y., Liu, C., Wei, S., Wei, C., and Liu, F., 2014. ECG quality assessment based on a kernel support vector machine and genetic algorithm with a feature matrix. *Journal of Zhejiang University*, 15(7), pp.564–573.

Zhou, H., Hou, K.-M., and Zuo, D., 2009. Real-Time Automatic ECG Diagnosis Method Dedicated to Pervasive Cardiac Care. *Wireless Sensor Network*, 1(4), pp.276–283.

Zidelmal, Z., Amirou, A., Adnane, M., and Belouchrani, A., 2012. QRS detection based on wavelet coefficients. *Computer Methods and Programs in Biomedicine*, 107(3), pp.490–496.

Appendices

Appendix I: Performance comparison between MMD and other QRS detection algorithms on MIT-BIH arrhythmia database

Signal	# beats	MMD				Pan and Tompkins (Pan & Tompkins 1985)				Wang et al. (Yeh & Wang 2008)				Li et al. (Pan et al. 2010)				JQRS (Johnson et al. 2015; Behar et al. 2014)			
		FP	FN	SN	PP	FP	FN	SN	PP	FP	FN	SN	PP	FP	FN	SN	PP	FP	FN	SN	PP
100	2273	0	2	99.91	100.00	0	0	100.00	100.00	0	1	99.96	100.00	0	0	100.00	100.00	0	0	100.00	100.00
101	1865	0	0	100.00	100.00	5	3	99.84	99.73	0	1	99.95	100.00	1	0	100.00	99.95	4	1	99.95	99.79
102	2187	0	0	100.00	100.00	0	0	100.00	100.00	0	1	99.95	100.00	0	0	100.00	100.00	0	1	99.95	100.00
103	2084	0	2	99.90	100.00	0	0	100.00	100.00	0	0	100.00	100.00	0	0	100.00	100.00	0	5	99.76	100.00
104	2230	3	3	99.87	99.87	1	0	100.00	99.96	2	0	100.00	99.91	8	2	99.91	99.64	13	23	98.97	99.41
105	2572	73	4	99.84	97.16	67	22	99.13	97.40	0	17	99.34	100.00	15	13	99.49	99.42	27	2	99.92	98.96
106	2027	0	19	99.07	100.00	5	2	99.90	99.75	0	6	99.70	100.00	2	3	99.85	99.90	0	65	96.79	100.00
107	2137	0	71	96.78	100.00	0	2	99.91	100.00	0	3	99.86	100.00	0	0	100.00	100.00	0	2	99.91	100.00
108	1763	13	3	99.83	99.26	199	22	98.61	88.71	6	0	100.00	99.66	13	15	99.15	99.26	27	7	99.60	98.49
109	2532	0	1	99.96	100.00	0	1	99.96	100.00	0	3	99.88	100.00	0	0	100.00	100.00	0	9	99.64	100.00
111	2124	0	1	99.95	100.00	1	0	100.00	99.95	0	1	99.95	100.00	1	1	99.95	99.95	0	2	99.91	100.00
112	2539	2	0	100.00	99.92	0	1	99.96	100.00	1	0	100.00	99.96	2	1	99.96	99.92	0	0	100.00	100.00
113	1795	0	2	99.89	100.00	0	0	100.00	100.00	9	0	100.00	99.50	2	0	100.00	99.89	0	0	100.00	100.00
114	1879	4	1	99.95	99.79	3	17	99.10	99.84	0	1	99.95	100.00	3	0	100.00	99.84	3	7	99.63	99.84
115	1953	0	2	99.90	100.00	0	0	100.00	100.00	0	0	100.00	100.00	0	0	100.00	100.00	0	0	100.00	100.00
116	2412	3	16	99.34	99.88	3	22	99.10	99.88	0	17	99.30	100.00	0	1	99.96	100.00	2	24	99.00	99.92
117	1535	1	0	100.00	99.93	1	1	99.93	99.93	2	0	100.00	99.87	1	0	100.00	99.93	0	0	100.00	100.00
118	2275	0	0	100.00	100.00	1	0	100.00	99.96	10	0	100.00	99.56	1	0	100.00	99.96	0	0	100.00	100.00
119	1987	0	0	100.00	100.00	1	0	100.00	99.95	0	0	100.00	100.00	1	0	100.00	99.95	0	2	99.90	100.00
121	1863	1	2	99.89	99.95	4	7	99.62	99.79	0	2	99.89	100.00	2	1	99.95	99.89	0	1	99.95	100.00
122	2476	0	0	100.00	100.00	1	1	99.96	99.96	0	0	100.00	100.00	0	0	100.00	100.00	0	1	99.96	100.00
123	1518	0	0	100.00	100.00	0	0	100.00	100.00	0	0	100.00	100.00	0	0	100.00	100.00	0	3	99.80	100.00
124	1619	0	0	100.00	100.00	0	0	100.00	100.00	1	0	100.00	99.94	0	0	100.00	100.00	0	11	99.32	100.00
200	2601	1	4	99.85	99.96	6	3	99.88	99.77	5	0	100.00	99.81	0	1	99.96	100.00	2	3	99.88	99.92
201	1963	0	51	97.47	100.00	0	10	99.49	100.00	0	20	98.99	100.00	1	12	99.39	99.95	0	67	96.59	100.00
202	2136	0	3	99.86	100.00	0	4	99.81	100.00	1	0	100.00	99.95	0	1	99.95	100.00	0	11	99.49	100.00
203	2982	12	66	97.83	99.60	53	30	98.99	98.22	16	2	99.93	99.46	2	24	99.20	99.93	10	60	97.99	99.66
205	2656	0	6	99.77	100.00	0	2	99.92	100.00	0	16	99.40	100.00	0	1	99.96	100.00	0	7	99.74	100.00

207	1862	224	10	99.39	87.97	4	4	99.79	99.79	0	1	99.95	100.00	2	3	99.84	99.89	303	57	96.94	85.61
208	2956	4	29	99.03	99.86	4	14	99.53	99.86	0	14	99.53	100.00	0	4	99.86	100.00	3	438	85.18	99.88
209	3004	3	0	100.00	99.90	3	0	100.00	99.90	1	0	100.00	99.97	0	0	100.00	100.00	1	0	100.00	99.97
210	2647	3	40	98.51	99.89	2	8	99.70	99.92	0	14	99.47	100.00	3	3	99.89	99.89	3	35	98.68	99.89
212	2748	0	0	100.00	100.00	0	0	100.00	100.00	1	0	100.00	99.96	0	0	100.00	100.00	0	0	100.00	100.00
213	3251	0	5	99.85	100.00	1	2	99.94	99.97	0	3	99.91	100.00	0	0	100.00	100.00	0	6	99.82	100.00
214	2262	0	4	99.82	100.00	2	4	99.82	99.91	0	4	99.82	100.00	–	–	–	–	0	10	99.56	100.00
215	3363	0	2	99.94	100.00	0	1	99.97	100.00	0	4	99.88	100.00	–	–	–	–	0	2	99.94	100.00
217	2208	0	6	99.73	100.00	4	6	99.73	99.82	0	2	99.91	100.00	1	1	99.95	99.95	0	6	99.73	100.00
219	2154	0	0	100.00	100.00	0	0	100.00	100.00	0	0	100.00	100.00	0	0	100.00	100.00	0	3	99.86	100.00
220	2048	0	0	100.00	100.00	0	0	100.00	100.00	0	0	100.00	100.00	0	0	100.00	100.00	0	0	100.00	100.00
221	2427	0	0	100.00	100.00	2	0	100.00	99.92	0	1	99.96	100.00	0	7	99.71	100.00	0	16	99.34	100.00
222	2484	1	6	99.76	99.96	101	81	96.71	95.93	0	5	99.80	100.00	1	9	99.64	99.96	1	6	99.76	99.96
223	2605	0	2	99.92	100.00	1	0	100.00	99.96	1	0	100.00	99.96	0	2	99.92	100.00	0	6	99.77	100.00
228	2053	19	19	99.07	99.07	25	5	99.75	98.78	0	2	99.90	100.00	3	7	99.66	99.85	11	4	99.81	99.47
230	2256	0	0	100.00	100.00	1	0	100.00	99.96	2	0	100.00	99.91	0	0	100.00	100.00	0	0	100.00	100.00
231	1886	0	0	100.00	100.00	0	0	100.00	100.00	0	15	99.21	100.00	0	0	100.00	100.00	0	0	100.00	100.00
232	1780	2	0	100.00	99.89	6	1	99.94	99.66	0	0	100.00	100.00	0	0	100.00	100.00	2	0	100.00	99.89
233	3079	0	7	99.77	100.00	0	1	99.97	100.00	0	9	99.71	100.00	0	0	100.00	100.00	0	16	99.48	100.00
234	2753	0	0	100.00	100.00	0	0	100.00	100.00	0	1	99.96	100.00	0	0	100.00	100.00	0	4	99.85	100.00

Appendix II: CEC Test suite 2014 results for 10 dimensions

	Min	SRFA	RFA	FA	ODFA	CFA	LSFA	SFA	SA	BSO	CSO	PSO	DFO	ALO
1	Mean	1.26E+00	1.75E+00	6.59E+05	6.23E+03	1.76E+07	2.68E+03	2.33E+03	6.48E+07	1.41E+08	9.33E+03	1.44E+05	6.14E+06	4.74E+08
	Min	3.03E-01	3.21E-01	4.17E+04	4.03E-01	5.28E+06	1.26E+01	5.25E+01	8.97E+06	1.29E+07	2.17E+03	5.09E+03	0.00E+00	5.30E+07
	Max	2.63E+00	3.35E+00	2.05E+06	3.97E+04	3.50E+07	6.47E+03	8.12E+03	1.12E+08	5.04E+08	2.13E+04	8.35E+05	3.51E+07	1.19E+09
	Std	5.76E-01	7.86E-01	5.03E+05	9.88E+03	8.81E+06	2.00E+03	2.15E+03	2.85E+07	1.17E+08	4.16E+03	2.22E+05	8.59E+06	2.84E+08
2	Mean	7.21E+01	1.34E+02	2.21E+03	4.19E+06	3.22E+09	3.84E+03	2.39E+03	7.23E+09	6.05E+09	7.38E+07	2.03E+03	3.14E+07	1.55E+10
	Min	2.86E+01	4.52E+01	1.52E+02	4.65E-02	1.51E+09	1.02E+01	9.67E-01	2.47E+09	1.66E+09	7.33E+03	2.94E+01	0.00E+00	6.09E+09
	Max	1.55E+02	2.62E+02	1.04E+04	3.12E+07	4.91E+09	1.03E+04	7.26E+03	1.04E+10	1.23E+10	2.09E+09	1.07E+04	5.17E+08	2.47E+10
	Std	2.68E+01	4.20E+01	2.68E+03	8.75E+06	8.80E+08	4.03E+03	2.32E+03	2.01E+09	2.99E+09	3.81E+08	2.48E+03	9.66E+07	4.56E+09
3	Mean	2.71E-03	2.96E-03	1.15E+04	5.95E+03	2.61E+04	4.70E+03	3.00E+03	7.75E+04	1.38E+05	6.27E-01	5.95E+03	2.66E+04	6.49E+07
	Min	2.46E-04	9.75E-04	3.79E+03	1.38E-01	1.51E+04	6.33E-02	7.59E-01	2.36E+04	1.74E+04	1.53E-01	3.49E+01	0.00E+00	5.65E+04
	Max	1.12E-02	4.88E-03	4.53E+04	4.60E+04	4.95E+04	1.56E+04	1.14E+04	4.39E+05	1.40E+06	2.22E+00	2.30E+04	1.00E+05	4.97E+08
	Std	2.69E-03	1.27E-03	7.80E+03	1.30E+04	7.93E+03	4.74E+03	3.62E+03	7.70E+04	2.70E+05	4.41E-01	5.99E+03	2.13E+04	1.15E+08
4	Mean	1.49E-05	2.46E-05	7.62E+00	4.03E-01	1.87E+02	9.80E-01	4.01E-01	4.86E+02	1.18E+03	2.06E+00	6.64E+00	2.52E+01	2.60E+03
	Min	5.04E-06	1.20E-05	6.06E+00	1.48E-09	5.32E+01	2.04E-07	2.11E-07	2.26E+02	1.32E+02	1.90E-01	9.28E-04	0.00E+00	3.00E+02
	Max	3.03E-05	4.74E-05	9.32E+00	3.83E+00	3.04E+02	4.02E+00	3.99E+00	8.74E+02	4.19E+03	4.76E+00	1.12E+02	1.51E+02	7.92E+03
	Std	5.48E-06	9.15E-06	8.20E-01	9.13E-01	5.72E+01	1.70E+00	1.22E+00	1.96E+02	1.05E+03	1.48E+00	2.01E+01	3.70E+01	1.64E+03
5	Mean	1.27E-02	1.82E-02	1.80E+01	2.83E+00	2.05E+01	2.00E+01	2.00E+01	2.07E+01	2.00E+01	1.99E+01	1.77E+01	1.92E+01	2.12E+01
	Min	8.90E-03	1.29E-02	2.06E-02	1.29E-04	2.03E+01	2.00E+01	2.00E+01	2.04E+01	2.00E+01	1.31E+01	3.40E+00	3.11E+00	2.08E+01
	Max	1.66E-02	2.28E-02	2.00E+01	1.13E+01	2.07E+01	2.00E+01	2.00E+01	2.11E+01	2.00E+01	2.05E+01	2.05E+01	2.08E+01	2.15E+01
	Std	1.89E-03	2.52E-03	6.10E+00	2.90E+00	1.14E-01	6.87E-04	1.04E-04	1.62E-01	2.32E-03	1.48E+00	4.93E+00	3.85E+00	1.88E-01
6	Mean	4.43E-02	5.66E-02	1.16E-01	6.12E-01	1.00E+01	1.35E+01	2.82E-01	1.16E+01	1.12E+01	4.95E+00	3.61E+00	4.65E+00	1.46E+01
	Min	3.05E-02	4.14E-02	6.03E-02	1.71E-03	8.07E+00	9.70E+00	1.17E-02	9.99E+00	8.27E+00	2.97E+00	7.78E-01	0.00E+00	1.11E+01
	Max	5.16E-02	6.67E-02	1.89E-01	3.42E+00	1.15E+01	1.55E+01	1.65E+00	1.39E+01	1.39E+01	6.75E+00	6.29E+00	1.26E+01	1.70E+01
	Std	4.84E-03	5.91E-03	3.39E-02	6.85E-01	6.83E-01	1.29E+00	3.49E-01	9.21E-01	1.31E+00	8.71E-01	1.63E+00	2.93E+00	1.49E+00
7	Mean	4.29E-04	9.03E-04	8.88E-02	5.97E-01	4.54E+01	5.32E+01	3.56E-01	9.74E+01	7.09E+01	7.66E-02	2.49E-01	1.59E+00	1.92E+02
	Min	2.23E-04	4.96E-04	6.78E-04	1.43E-07	2.62E+01	3.04E+01	1.97E-02	4.78E+01	1.94E+01	4.89E-02	3.10E-02	0.00E+00	8.26E+01
	Max	7.32E-04	1.73E-03	8.35E-01	1.41E+00	7.03E+01	8.92E+01	9.23E-01	1.37E+02	2.00E+02	1.26E-01	1.01E+00	4.17E+00	2.65E+02
	Std	1.19E-04	2.69E-04	2.00E-01	4.12E-01	1.01E+01	1.30E+01	2.22E-01	2.14E+01	4.00E+01	1.97E-02	2.53E-01	1.08E+00	4.34E+01
8	Mean	3.44E-05	7.12E-05	1.04E+01	7.54E-01	6.63E+01	5.36E+01	2.99E+01	8.89E+01	8.34E+01	1.23E+01	1.51E+01	2.82E+01	1.24E+02
	Min	7.23E-06	3.18E-05	3.98E+00	6.49E-09	5.29E+01	2.29E+01	5.97E+00	6.97E+01	5.40E+01	8.06E+00	6.97E+00	4.17E+00	9.23E+01
	Max	5.42E-05	1.24E-04	3.68E+01	8.96E+00	7.92E+01	8.36E+01	4.88E+01	1.05E+02	1.18E+02	1.78E+01	2.89E+01	5.41E+01	1.55E+02
	Std	1.13E-05	2.35E-05	6.58E+00	2.02E+00	7.72E+00	1.57E+01	1.20E+01	1.03E+01	1.46E+01	2.73E+00	5.44E+00	1.39E+01	1.54E+01
9	Mean	5.01E-05	8.60E-05	9.88E+00	2.27E+00	7.08E+01	5.98E+01	2.86E+01	9.47E+01	9.31E+01	2.21E+01	1.73E+01	3.01E+01	1.35E+02
	Min	1.83E-05	2.31E-05	2.98E+00	3.72E-08	5.46E+01	2.69E+01	5.97E+00	7.13E+01	4.96E+01	1.05E+01	2.99E+00	0.00E+00	8.24E+01
	Max	9.11E-05	1.73E-04	2.69E+01	2.99E+01	8.75E+01	1.11E+02	5.47E+01	1.20E+02	1.39E+02	3.35E+01	4.38E+01	7.51E+01	1.74E+02
	Std	1.81E-05	3.11E-05	5.31E+00	5.96E+00	8.26E+00	1.93E+01	1.33E+01	1.12E+01	1.78E+01	6.25E+00	8.06E+00	2.11E+01	1.98E+01
10	Mean	9.85E-04	1.54E-03	3.73E+02	3.19E+00	1.71E+03	1.44E+03	1.01E+03	2.09E+03	2.05E+03	4.07E+02	6.66E+02	4.48E+02	2.86E+03
	Min	5.19E-04	6.25E-04	3.54E+00	1.34E-03	1.34E+03	3.39E+02	2.45E+02	1.54E+03	8.35E+02	1.45E+02	3.60E+00	0.00E+00	2.18E+03
	Max	1.43E-03	2.56E-03	9.75E+02	3.62E+01	1.98E+03	2.10E+03	1.84E+03	2.39E+03	2.78E+03	6.54E+02	1.48E+03	1.59E+03	3.33E+03
	Std	2.47E-04	5.38E-04	2.86E+02	7.39E+00	1.54E+02	4.52E+02	3.74E+02	1.87E+02	3.86E+02	1.34E+02	3.39E+02	4.15E+02	3.20E+02
11	Mean	1.24E-03	2.19E-03	4.50E+02	4.58E+01	1.57E+03	1.36E+03	9.29E+02	1.95E+03	2.01E+03	6.97E+02	8.03E+02	1.25E+03	2.83E+03
	Min	4.50E-04	7.66E-04	1.36E+01	2.60E-10	1.32E+03	4.31E+02	2.39E+02	1.38E+03	1.08E+03	4.00E+02	2.24E+02	2.16E+00	1.81E+03

	Max	2.27E-03	3.85E-03	1.28E+03	9.23E+02	1.91E+03	2.31E+03	1.56E+03	2.37E+03	2.96E+03	1.01E+03	1.75E+03	2.20E+03	3.20E+03
	Std	4.36E-04	8.55E-04	3.02E+02	1.69E+02	1.67E+02	4.20E+02	3.67E+02	2.55E+02	4.25E+02	1.28E+02	3.41E+02	5.63E+02	3.36E+02
12	Mean	4.52E-03	6.03E-03	4.45E-02	2.78E-01	1.52E+00	8.69E-01	3.47E-01	2.71E+00	1.71E+00	3.88E-01	5.26E-01	1.00E+00	5.99E+00
	Min	3.44E-03	3.56E-03	9.03E-03	1.93E-04	8.04E-01	1.56E-02	2.43E-04	1.45E+00	9.84E-01	2.67E-01	1.18E-01	1.51E-01	3.43E+00
	Max	5.74E-03	8.08E-03	2.02E-01	1.17E+00	2.22E+00	3.55E+00	1.07E+00	5.58E+00	2.98E+00	5.06E-01	1.05E+00	2.39E+00	9.01E+00
	Std	5.73E-04	9.54E-04	4.65E-02	3.32E-01	3.68E-01	8.35E-01	2.85E-01	9.77E-01	6.22E-01	7.18E-02	2.60E-01	5.24E-01	1.49E+00
13	Mean	4.27E-02	3.62E-02	1.16E-01	2.50E-01	2.11E+00	5.32E-01	2.89E-01	2.83E+00	2.94E+00	2.37E-01	3.30E-01	3.36E-01	5.47E+00
	Min	1.38E-02	1.02E-02	6.75E-02	4.36E-02	1.58E+00	1.98E-01	1.64E-01	1.20E+00	7.78E-01	1.28E-01	1.09E-01	0.00E+00	3.51E+00
	Max	6.57E-02	5.35E-02	2.21E-01	5.40E-01	2.65E+00	9.08E-01	5.05E-01	4.39E+00	4.33E+00	3.78E-01	6.26E-01	9.77E-01	6.91E+00
	Std	1.24E-02	1.14E-02	3.75E-02	1.17E-01	3.14E-01	2.00E-01	9.33E-02	6.47E-01	9.23E-01	5.51E-02	1.25E-01	2.24E-01	8.07E-01
14	Mean	5.82E-03	4.27E-03	3.58E-01	1.39E-01	1.20E+01	3.45E-01	3.04E-01	2.50E+01	2.24E+01	1.98E-01	3.03E-01	2.48E-01	5.41E+01
	Min	4.89E-04	7.22E-04	2.02E-01	4.07E-02	6.10E+00	7.66E-02	1.49E-01	5.39E+00	4.65E+00	1.17E-01	9.32E-02	0.00E+00	2.00E+01
	Max	1.46E-02	1.01E-02	4.77E-01	2.96E-01	2.11E+01	1.23E+00	4.74E-01	3.73E+01	5.73E+01	3.00E-01	1.20E+00	9.24E-01	7.59E+01
	Std	4.43E-03	2.24E-03	7.55E-02	6.32E-02	3.60E+00	3.38E-01	9.20E-02	6.79E+00	1.16E+01	4.73E-02	2.10E-01	1.80E-01	1.40E+01
15	Mean	1.59E-09	4.81E-09	1.15E+00	1.00E-01	1.26E+03	5.48E+02	1.85E+00	1.73E+04	3.84E+04	1.65E+00	2.02E+00	5.87E+00	3.95E+05
	Min	4.60E-10	2.20E-10	6.65E-01	0.00E+00	9.54E+01	1.99E+02	6.53E-01	3.26E+03	3.35E+02	1.11E+00	5.45E-01	2.40E-02	2.68E+03
	Max	4.16E-09	2.01E-08	1.94E+00	1.66E+00	3.48E+03	8.14E+02	5.51E+00	7.70E+04	1.35E+05	2.51E+00	4.64E+00	4.18E+01	1.72E+06
	Std	8.67E-10	4.68E-09	3.21E-01	3.49E-01	8.56E+02	1.55E+02	1.04E+00	1.53E+04	4.18E+04	3.67E-01	9.40E-01	7.70E+00	3.96E+05
16	Mean	1.89E-04	3.46E-04	3.11E+00	5.42E-01	3.73E+00	4.15E+00	3.44E+00	4.09E+00	4.07E+00	3.33E+00	2.81E+00	2.82E+00	4.60E+00
	Min	8.62E-05	1.71E-04	1.55E+00	1.45E-08	2.83E+00	3.00E+00	2.67E+00	3.66E+00	3.55E+00	2.87E+00	1.79E+00	0.00E+00	4.06E+00
	Max	2.78E-04	7.42E-04	3.76E+00	2.21E+00	3.97E+00	4.82E+00	4.04E+00	4.39E+00	4.67E+00	3.58E+00	3.65E+00	3.92E+00	4.90E+00
	Std	5.09E-05	1.37E-04	4.37E-01	6.70E-01	2.41E-01	3.83E-01	3.78E-01	1.76E-01	3.06E-01	1.84E-01	4.88E-01	1.05E+00	1.83E-01
17	Mean	3.95E-02	5.48E-02	2.33E+04	3.45E+03	2.57E+05	4.84E+03	2.84E+03	1.56E+06	7.76E+06	1.56E+02	7.53E+03	2.23E+05	4.92E+07
	Min	3.27E-03	1.80E-02	4.87E+02	1.49E-01	1.75E+04	4.91E+02	2.71E+02	1.12E+05	1.08E+04	4.59E+01	7.83E+02	0.00E+00	1.62E+06
	Max	1.17E-01	2.00E-01	1.88E+05	7.92E+04	9.63E+05	1.54E+04	1.19E+04	6.10E+06	6.14E+07	3.07E+02	5.68E+04	2.11E+06	1.75E+08
	Std	3.26E-02	4.13E-02	4.49E+04	1.45E+04	1.99E+05	4.54E+03	3.24E+03	1.26E+06	1.32E+07	6.24E+01	1.10E+04	5.07E+05	5.07E+07
18	Mean	3.60E-01	3.34E-01	3.09E+03	1.63E+05	1.39E+06	3.70E+03	2.23E+03	1.33E+07	4.14E+07	8.45E+00	2.51E+03	1.89E+03	5.34E+08
	Min	5.38E-02	7.09E-02	9.97E+01	6.63E-02	5.65E+02	1.52E+01	1.29E+01	3.61E+05	1.97E+02	4.67E+00	1.84E+01	0.00E+00	3.00E+07
	Max	1.20E+00	1.06E+00	1.14E+04	1.69E+06	4.58E+06	1.26E+04	1.21E+04	8.44E+07	3.23E+08	1.55E+01	1.15E+04	1.35E+04	1.75E+09
	Std	3.25E-01	2.76E-01	3.18E+03	4.04E+05	1.23E+06	3.81E+03	3.05E+03	1.69E+07	6.82E+07	2.95E+00	3.23E+03	4.51E+03	4.56E+08
19	Mean	4.34E-03	5.66E-03	2.67E+00	7.80E-01	1.05E+01	4.95E+01	2.78E+00	2.34E+01	5.39E+01	2.35E+00	3.55E+00	3.14E+00	9.49E+01
	Min	1.52E-03	3.65E-03	1.22E+00	1.07E-03	7.20E+00	1.97E+01	1.20E+00	1.18E+01	1.42E+01	1.85E+00	1.54E+00	0.00E+00	3.28E+01
	Max	6.07E-03	9.36E-03	5.77E+00	4.62E+00	1.38E+01	1.32E+02	5.12E+00	4.22E+01	1.55E+02	3.47E+00	6.76E+00	9.61E+00	3.05E+02
	Std	1.19E-03	1.53E-03	8.68E-01	1.17E+00	2.08E+00	2.67E+01	9.74E-01	8.53E+00	3.76E+01	3.68E-01	1.37E+00	2.56E+00	5.47E+01
20	Mean	3.67E-02	3.40E-02	2.62E+03	3.00E+03	1.42E+04	4.46E+03	2.06E+03	1.75E+05	5.02E+06	4.34E+00	2.02E+03	5.31E+04	1.42E+08
	Min	1.32E-02	7.17E-03	8.16E+01	1.42E+00	8.70E+02	1.43E+02	6.04E+01	8.10E+03	7.27E+02	2.17E+00	3.14E+01	0.00E+00	1.52E+05
	Max	8.62E-02	7.34E-02	9.78E+03	1.87E+04	4.22E+04	1.24E+04	9.87E+03	8.71E+05	3.84E+07	8.81E+00	1.31E+04	4.05E+05	1.07E+09
	Std	1.62E-02	1.57E-02	2.58E+03	5.22E+03	9.08E+03	4.08E+03	2.34E+03	1.92E+05	9.61E+06	1.47E+00	3.02E+03	9.42E+04	2.62E+08
21	Mean	5.17E-02	5.96E-02	3.38E+03	4.50E+03	3.98E+04	2.84E+03	2.06E+03	4.98E+05	4.75E+06	3.63E+01	1.66E+03	2.87E+05	4.06E+07
	Min	1.34E-02	2.17E-02	4.44E+02	2.75E+00	3.02E+03	3.15E+02	1.61E+02	1.62E+04	8.09E+03	7.21E+00	1.87E+02	0.00E+00	3.53E+05
	Max	1.11E-01	1.07E-01	1.10E+04	7.33E+04	1.42E+05	8.48E+03	6.10E+03	2.49E+06	5.00E+07	8.05E+01	7.90E+03	5.40E+06	3.52E+08
	Std	2.30E-02	2.06E-02	2.94E+03	1.35E+04	3.23E+04	2.78E+03	1.57E+03	5.76E+05	1.09E+07	1.94E+01	1.91E+03	1.12E+06	6.83E+07
22	Mean	1.28E-01	1.59E-01	3.29E+01	5.70E+00	1.44E+02	3.87E+02	3.51E+01	3.27E+02	4.74E+02	2.28E+01	4.38E+01	4.22E+01	9.24E+02
	Min	7.23E-02	6.92E-02	2.28E+01	5.34E-02	6.71E+01	7.35E+01	2.01E+01	1.23E+02	1.20E+02	1.49E+01	2.02E+01	0.00E+00	4.08E+02
	Max	1.80E-01	2.13E-01	8.74E+01	3.96E+01	2.39E+02	7.99E+02	1.43E+02	4.97E+02	1.28E+03	2.59E+01	1.43E+02	4.49E+02	1.95E+03
	Std	2.37E-02	3.09E-02	1.26E+01	8.19E+00	4.46E+01	1.72E+02	2.45E+01	1.05E+02	2.36E+02	1.76E+00	3.04E+01	8.56E+01	3.73E+02

23	Mean	2.00E+02	2.00E+02	2.00E+02	2.01E+02	2.94E+02	1.99E+02	2.00E+02	3.54E+02	4.11E+02	1.93E+02	2.00E+02	2.16E+02	7.56E+02
	Min	2.00E+02	2.00E+02	2.00E+02	2.00E+02	2.52E+02	1.86E+02	1.86E+02	2.90E+02	2.68E+02	1.86E+02	2.00E+02	2.00E+02	4.42E+02
	Max	2.00E+02	2.00E+02	2.00E+02	2.09E+02	3.20E+02	2.00E+02	2.00E+02	4.20E+02	6.01E+02	2.00E+02	2.00E+02	2.96E+02	1.24E+03
	Std	1.34E-03	2.26E-03	3.85E-03	1.66E+00	1.47E+01	3.50E+00	2.52E+00	3.34E+01	8.77E+01	6.95E+00	1.85E-02	2.19E+01	2.02E+02
24	Mean	1.99E+02	1.98E+02	1.77E+02	4.68E+01	2.12E+02	2.29E+02	1.83E+02	2.29E+02	2.33E+02	1.67E+02	1.90E+02	1.97E+02	2.70E+02
	Min	1.95E+02	1.73E+02	1.51E+02	3.51E+01	1.96E+02	1.96E+02	1.52E+02	2.08E+02	2.08E+02	1.52E+02	1.59E+02	1.65E+02	2.41E+02
	Max	2.00E+02	1.99E+02	2.00E+02	1.37E+02	2.24E+02	2.68E+02	2.05E+02	2.45E+02	2.74E+02	1.91E+02	2.04E+02	2.25E+02	2.97E+02
	Std	8.15E-01	5.11E+00	1.26E+01	2.78E+01	7.37E+00	1.67E+01	1.47E+01	7.81E+00	1.52E+01	8.57E+00	1.38E+01	1.08E+01	1.47E+01
25	Mean	1.98E+02	1.96E+02	1.94E+02	1.96E+02	2.03E+02	2.03E+02	1.95E+02	2.08E+02	2.12E+02	1.90E+02	1.93E+02	1.97E+02	2.34E+02
	Min	1.96E+02	1.93E+02	1.80E+02	1.88E+02	2.00E+02	1.95E+02	1.79E+02	2.03E+02	2.02E+02	1.74E+02	1.70E+02	1.81E+02	2.10E+02
	Max	1.98E+02	1.99E+02	2.00E+02	2.00E+02	2.06E+02	2.27E+02	2.00E+02	2.14E+02	2.29E+02	1.98E+02	2.00E+02	2.08E+02	2.60E+02
	Std	6.65E-01	1.66E+00	4.81E+00	3.17E+00	1.33E+00	5.92E+00	5.47E+00	2.71E+00	6.66E+00	4.68E+00	6.80E+00	6.04E+00	1.64E+01
26	Mean	1.05E+02	1.04E+02	1.02E+02	1.18E+02	1.04E+02	1.12E+02	1.02E+02	1.06E+02	1.17E+02	1.02E+02	1.19E+02	1.11E+02	1.52E+02
	Min	1.04E+02	1.03E+02	1.02E+02	1.05E+02	1.04E+02	1.02E+02	1.02E+02	1.04E+02	1.06E+02	1.02E+02	1.02E+02	1.02E+02	1.06E+02
	Max	1.06E+02	1.04E+02	1.02E+02	2.00E+02	1.05E+02	2.00E+02	1.02E+02	1.08E+02	2.08E+02	1.02E+02	2.00E+02	2.05E+02	2.52E+02
	Std	4.43E-01	2.80E-01	5.36E-02	2.80E+01	3.05E-01	2.98E+01	1.06E-01	1.14E+00	2.08E+01	4.07E-02	3.69E+01	2.51E+01	5.66E+01
27	Mean	2.00E+02	4.53E+02	3.97E+02	2.30E+02	4.85E+02	6.52E+02	4.04E+02	5.18E+02	6.20E+02	4.05E+02	4.50E+02	5.07E+02	7.52E+02
	Min	2.00E+02	4.16E+02	3.68E+02	2.00E+02	4.25E+02	4.04E+02	3.72E+02	4.49E+02	4.81E+02	4.02E+02	3.93E+02	2.17E+02	6.53E+02
	Max	2.00E+02	5.67E+02	4.43E+02	4.30E+02	5.47E+02	1.08E+03	4.51E+02	5.99E+02	7.03E+02	4.11E+02	5.52E+02	6.71E+02	7.96E+02
	Std	1.07E-03	3.64E+01	2.15E+01	5.53E+01	3.29E+01	1.59E+02	1.84E+01	3.65E+01	5.16E+01	2.60E+00	4.90E+01	9.33E+01	4.50E+01
28	Mean	2.00E+02	5.24E+02	3.91E+02	2.36E+02	4.23E+02	1.27E+03	4.64E+02	1.13E+03	9.59E+02	3.62E+02	5.97E+02	7.02E+02	2.26E+03
	Min	2.00E+02	3.97E+02	3.49E+02	2.00E+02	3.98E+02	7.66E+02	3.49E+02	8.16E+02	3.72E+02	3.52E+02	3.93E+02	4.17E+02	1.13E+03
	Max	2.00E+02	8.87E+02	6.08E+02	8.19E+02	4.42E+02	2.12E+03	9.46E+02	1.54E+03	2.13E+03	3.85E+02	1.20E+03	1.10E+03	3.18E+03
	Std	1.92E-03	1.06E+02	7.16E+01	1.16E+02	9.50E+00	3.66E+02	1.65E+02	1.88E+02	4.58E+02	1.02E+01	2.40E+02	1.89E+02	4.90E+02
29	Mean	2.00E+02	2.00E+02	8.59E+02	1.56E+03	6.05E+03	4.34E+02	3.69E+02	2.21E+05	4.31E+04	2.18E+02	6.35E+02	1.75E+05	8.97E+06
	Min	2.00E+02	2.00E+02	2.89E+02	2.00E+02	1.33E+03	2.71E+02	2.51E+02	1.69E+04	8.72E+02	2.10E+02	2.55E+02	1.99E+02	1.38E+05
	Max	2.00E+02	2.00E+02	3.22E+03	1.67E+04	2.65E+04	5.73E+02	7.27E+02	8.79E+05	2.75E+05	2.47E+02	2.82E+03	1.80E+06	2.58E+07
	Std	1.00E-03	2.17E-03	7.05E+02	3.71E+03	5.26E+03	7.48E+01	1.19E+02	2.02E+05	6.31E+04	7.70E+00	5.21E+02	4.35E+05	7.28E+06
30	Mean	2.00E+02	2.00E+02	5.77E+02	2.36E+02	2.39E+03	1.46E+03	7.10E+02	3.26E+04	1.49E+05	3.21E+02	7.87E+02	3.85E+03	1.09E+06
	Min	2.00E+02	2.00E+02	3.05E+02	2.00E+02	1.13E+03	4.02E+02	2.91E+02	3.61E+03	8.84E+02	2.82E+02	3.61E+02	4.89E+02	1.08E+05
	Max	2.00E+02	2.00E+02	1.05E+03	6.78E+02	5.61E+03	1.42E+04	9.96E+02	1.06E+05	1.99E+06	4.49E+02	1.90E+03	1.99E+04	3.51E+06
	Std	2.54E-02	2.03E-02	1.79E+02	8.87E+01	1.09E+03	2.44E+03	2.17E+02	2.72E+04	3.93E+05	3.27E+01	3.45E+02	4.30E+03	9.88E+05

Appendix III: CEC Test suite 2014 results for 30 dimensions

		SRFA	RFA	FA	ODFA	CFA	LSFA	SFA	SA	BSO	CSO	PSO	DFO	ALO
1	Mean	1.93E+01	3.98E+01	1.29E+07	5.53E+06	7.81E+08	8.96E+05	9.70E+04	1.33E+09	1.59E+09	1.02E+07	8.20E+06	1.37E+08	3.50E+09
	Min	1.05E+01	1.85E+01	3.49E+06	3.04E+00	3.98E+08	5.41E+04	1.60E+04	6.85E+08	6.14E+08	3.20E+06	1.44E+06	1.35E+06	1.18E+09
	Max	2.94E+01	6.37E+01	2.57E+07	1.40E+08	9.98E+08	6.62E+06	3.27E+05	1.82E+09	3.39E+09	2.06E+07	5.18E+07	6.72E+08	7.12E+09
	Std	5.24E+00	1.22E+01	6.98E+06	2.56E+07	1.43E+08	1.35E+06	8.11E+04	2.78E+08	7.09E+08	3.45E+06	1.00E+07	1.56E+08	1.31E+09
2	Mean	7.51E+02	2.09E+03	3.09E+04	9.97E+07	5.71E+10	3.77E+03	3.01E+03	7.69E+10	4.45E+10	1.00E+10	6.71E+07	1.71E+09	1.10E+11
	Min	5.05E+02	1.09E+03	1.24E+04	1.68E-01	4.49E+10	8.84E-01	5.08E+01	5.93E+10	2.11E+10	1.00E+10	7.07E+05	0.00E+00	6.17E+10
	Max	1.01E+03	2.72E+03	4.37E+04	2.54E+09	6.54E+10	1.06E+04	1.05E+04	9.08E+10	9.22E+10	1.00E+10	7.73E+08	9.77E+09	1.37E+11
	Std	1.08E+02	3.77E+02	8.86E+03	4.63E+08	5.60E+09	4.07E+03	3.50E+03	7.88E+09	1.93E+10	0.00E+00	1.43E+08	2.55E+09	1.65E+10
3	Mean	6.10E-03	1.08E-02	6.15E+04	2.81E+03	1.24E+05	1.59E+04	4.86E+03	2.02E+05	7.61E+05	1.28E+03	2.13E+04	8.50E+04	5.02E+07
	Min	1.77E-03	4.19E-03	3.86E+04	3.38E-03	7.85E+04	2.22E+03	2.01E+02	1.06E+05	3.15E+04	4.98E+02	1.36E+03	1.33E+02	1.61E+05
	Max	1.98E-02	1.96E-02	1.36E+05	3.41E+04	1.70E+05	5.60E+04	2.02E+04	4.46E+05	7.05E+06	2.24E+03	7.05E+04	2.40E+05	2.95E+08
	Std	3.84E-03	3.91E-03	1.88E+04	6.45E+03	2.03E+04	1.29E+04	4.83E+03	7.24E+04	1.42E+06	4.84E+02	1.44E+04	6.74E+04	7.97E+07
4	Mean	1.34E-04	3.79E-04	3.00E+01	3.97E+00	5.90E+03	1.14E+02	3.14E+01	1.29E+04	9.20E+03	7.63E+01	1.17E+02	3.79E+02	2.61E+04
	Min	7.40E-05	2.43E-04	2.47E+01	1.31E-04	3.48E+03	1.77E+01	2.06E+01	8.48E+03	2.96E+03	2.43E+01	1.87E+01	7.71E+00	1.58E+04
	Max	1.90E-04	5.11E-04	8.48E+01	5.70E+01	8.22E+03	2.27E+02	8.22E+01	1.70E+04	2.61E+04	1.50E+02	2.08E+02	1.09E+03	4.29E+04
	Std	2.49E-05	6.62E-05	1.04E+01	1.09E+01	1.07E+03	4.38E+01	1.86E+01	2.06E+03	5.61E+03	2.90E+01	4.63E+01	2.70E+02	8.04E+03
5	Mean	2.04E-02	3.44E-02	2.00E+01	2.75E+00	2.11E+01	2.00E+01	2.00E+01	2.12E+01	2.00E+01	2.10E+01	2.07E+01	2.10E+01	2.15E+01
	Min	1.65E-02	2.62E-02	2.00E+01	3.09E-04	2.08E+01	2.00E+01	2.00E+01	2.11E+01	2.00E+01	2.07E+01	2.00E+01	2.07E+01	2.12E+01
	Max	2.36E-02	4.23E-02	2.01E+01	1.17E+01	2.12E+01	2.00E+01	2.00E+01	2.13E+01	2.00E+01	2.11E+01	2.12E+01	2.13E+01	2.16E+01
	Std	1.56E-03	3.85E-03	1.87E-02	3.52E+00	6.61E-02	6.09E-03	4.30E-05	7.08E-02	5.65E-05	8.46E-02	3.39E-01	1.69E-01	8.32E-02
6	Mean	1.77E-01	2.49E-01	2.02E+00	2.12E+00	4.16E+01	4.85E+01	4.77E+00	4.46E+01	4.32E+01	3.07E+01	2.27E+01	2.12E+01	5.07E+01
	Min	1.63E-01	2.22E-01	6.87E-01	5.44E-03	3.81E+01	4.40E+01	2.32E+00	4.11E+01	3.77E+01	2.66E+01	1.84E+01	0.00E+00	4.49E+01
	Max	1.98E-01	2.86E-01	5.33E+00	9.64E+00	4.35E+01	5.18E+01	9.27E+00	4.65E+01	4.82E+01	3.52E+01	2.79E+01	4.16E+01	5.32E+01
	Std	9.45E-03	1.37E-02	1.17E+00	2.71E+00	1.35E+00	2.14E+00	1.70E+00	1.48E+00	3.07E+00	1.71E+00	2.23E+00	9.82E+00	1.89E+00
7	Mean	1.31E-03	3.74E-03	1.08E-01	7.28E-01	5.10E+02	8.20E+00	3.31E-01	6.58E+02	3.38E+02	1.15E+00	2.79E+00	1.43E+01	9.33E+02
	Min	7.54E-04	2.43E-03	6.71E-02	4.95E-08	3.53E+02	7.28E-01	6.39E-05	4.15E+02	1.68E+02	1.03E+00	1.18E+00	0.00E+00	6.06E+02
	Max	1.79E-03	5.72E-03	2.13E-01	2.84E+00	5.84E+02	2.27E+01	9.87E-01	7.91E+02	6.10E+02	1.27E+00	1.50E+01	9.19E+01	1.15E+03
	Std	2.66E-04	9.08E-04	3.28E-02	7.35E-01	5.44E+01	6.31E+00	2.62E-01	7.43E+01	1.08E+02	7.20E-02	2.54E+00	2.04E+01	1.28E+02
8	Mean	2.66E-04	7.08E-04	4.49E+01	3.98E+00	3.35E+02	2.05E+02	1.25E+02	3.82E+02	3.17E+02	9.56E+01	7.31E+01	1.33E+02	4.46E+02
	Min	1.91E-04	4.80E-04	1.49E+01	8.15E-09	2.82E+02	1.42E+02	7.56E+01	3.40E+02	2.38E+02	7.22E+01	4.21E+01	3.06E+01	3.93E+02
	Max	3.24E-04	9.51E-04	9.55E+01	2.90E+01	3.70E+02	2.78E+02	2.06E+02	4.15E+02	3.93E+02	1.23E+02	1.27E+02	2.47E+02	4.93E+02
	Std	3.27E-05	1.25E-04	1.73E+01	7.87E+00	1.95E+01	3.32E+01	3.21E+01	2.30E+01	3.94E+01	1.31E+01	2.02E+01	6.41E+01	2.43E+01
9	Mean	4.03E-04	1.14E-03	4.20E+01	1.20E+01	4.05E+02	2.81E+02	1.72E+02	4.67E+02	3.65E+02	1.36E+02	8.81E+01	1.41E+02	5.84E+02
	Min	2.96E-04	8.29E-04	1.69E+01	1.11E-07	3.54E+02	1.81E+02	1.02E+02	3.96E+02	2.94E+02	1.02E+02	4.27E+01	0.00E+00	5.09E+02
	Max	5.24E-04	1.67E-03	8.06E+01	1.80E+02	4.53E+02	4.29E+02	2.45E+02	5.17E+02	4.99E+02	1.67E+02	1.34E+02	2.30E+02	6.61E+02
	Std	6.62E-05	2.06E-04	1.39E+01	3.58E+01	2.45E+01	5.09E+01	3.39E+01	2.87E+01	4.38E+01	1.64E+01	2.31E+01	6.19E+01	3.64E+01
10	Mean	6.16E-03	1.65E-02	2.34E+03	1.09E+01	8.13E+03	4.89E+03	4.26E+03	8.74E+03	8.58E+03	3.70E+03	3.22E+03	2.11E+03	1.02E+04
	Min	4.38E-03	9.73E-03	7.97E+02	2.08E-05	6.94E+03	3.27E+03	2.81E+03	7.89E+03	5.44E+03	3.00E+03	1.89E+03	0.00E+00	9.31E+03
	Max	8.55E-03	2.22E-02	3.92E+03	3.31E+01	8.78E+03	6.74E+03	5.87E+03	9.79E+03	1.03E+04	4.38E+03	4.81E+03	5.42E+03	1.09E+04
	Std	9.76E-04	3.12E-03	7.75E+02	1.26E+01	3.53E+02	8.16E+02	7.08E+02	4.51E+02	1.08E+03	3.91E+02	7.68E+02	1.82E+03	3.99E+02
11	Mean	9.38E-03	2.62E-02	2.43E+03	9.75E+01	7.70E+03	4.60E+03	4.31E+03	8.46E+03	8.36E+03	4.43E+03	3.62E+03	6.08E+03	1.00E+04
	Min	4.89E-03	1.81E-02	7.92E+02	1.10E-05	6.77E+03	3.49E+03	2.17E+03	7.51E+03	6.30E+03	3.48E+03	2.24E+03	9.43E+02	9.22E+03

	Max	1.20E-02	3.29E-02	3.85E+03	7.18E+02	8.23E+03	5.72E+03	5.66E+03	8.98E+03	9.45E+03	5.08E+03	5.73E+03	7.74E+03	1.09E+04
	Std	1.55E-03	3.90E-03	6.42E+02	1.70E+02	3.56E+02	6.11E+02	7.08E+02	4.20E+02	8.69E+02	3.38E+02	8.94E+02	1.41E+03	4.54E+02
12	Mean	2.02E-03	2.32E-01	6.88E-02	2.71E-01	3.61E+00	7.81E-01	7.18E-01	4.34E+00	3.29E+00	1.53E+00	1.05E+00	2.42E+00	8.17E+00
	Min	1.41E-03	2.62E-03	2.57E-02	3.19E-05	2.52E+00	1.99E-01	2.02E-01	2.81E+00	1.68E+00	9.68E-01	3.77E-01	4.85E-02	4.48E+00
	Max	2.38E-03	3.71E+00	1.48E-01	1.52E+00	4.58E+00	1.95E+00	1.91E+00	5.43E+00	5.04E+00	1.96E+00	1.79E+00	5.33E+00	1.11E+01
	Std	2.04E-04	8.73E-01	3.17E-02	4.37E-01	5.03E-01	4.01E-01	4.20E-01	7.33E-01	8.73E-01	2.56E-01	3.89E-01	1.41E+00	1.58E+00
13	Mean	5.20E-02	4.35E-02	3.79E-01	2.17E-01	5.82E+00	9.22E-01	6.18E-01	7.02E+00	4.81E+00	3.58E-01	5.38E-01	3.91E-01	8.57E+00
	Min	1.99E-02	2.11E-02	2.37E-01	5.44E-02	4.96E+00	3.87E-01	2.48E-01	5.93E+00	3.20E+00	2.27E-01	3.20E-01	0.00E+00	6.82E+00
	Max	8.33E-02	6.84E-02	5.37E-01	5.04E-01	6.61E+00	2.86E+00	9.34E-01	8.60E+00	7.12E+00	5.87E-01	8.21E-01	2.10E+00	1.04E+01
	Std	1.67E-02	1.10E-02	6.73E-02	8.59E-02	3.83E-01	6.44E-01	1.67E-01	6.44E-01	8.48E-01	8.25E-02	1.23E-01	4.87E-01	8.39E-01
14	Mean	1.35E-02	1.32E-02	4.70E-01	1.56E+00	1.58E+02	4.21E-01	5.38E-01	2.19E+02	1.43E+02	2.67E-01	2.89E-01	6.96E+00	3.28E+02
	Min	4.32E-03	1.34E-03	2.71E-01	1.25E-03	1.37E+02	1.38E-01	2.91E-01	1.74E+02	6.96E+01	1.43E-01	2.17E-01	3.24E-02	2.50E+02
	Max	2.65E-02	2.56E-02	9.05E-01	2.54E+01	1.87E+02	6.06E+00	1.04E+00	2.43E+02	3.00E+02	3.58E-01	3.92E-01	2.67E+01	4.10E+02
	Std	5.69E-03	6.70E-03	1.80E-01	4.66E+00	1.41E+01	1.07E+00	2.56E-01	1.59E+01	4.37E+01	4.46E-02	4.51E-02	8.83E+00	3.62E+01
15	Mean	3.69E-08	2.78E-07	4.35E+00	2.72E+00	1.41E+06	3.64E+03	1.13E+01	3.60E+06	7.50E+05	2.06E+01	9.88E+01	2.61E+02	2.08E+07
	Min	1.34E-08	1.64E-07	2.83E+00	0.00E+00	4.38E+05	2.52E+03	4.27E+00	1.55E+06	2.77E+04	1.58E+01	3.34E+01	2.64E+01	2.47E+06
	Max	7.87E-08	4.62E-07	6.43E+00	2.82E+01	3.38E+06	4.38E+03	2.49E+01	5.42E+06	4.79E+06	3.06E+01	4.11E+02	2.05E+03	4.83E+07
	Std	1.42E-08	9.31E-08	9.48E-01	7.06E+00	6.45E+05	5.43E+02	5.83E+00	9.84E+05	1.13E+06	3.57E+00	7.84E+01	4.39E+02	1.21E+07
16	Mean	1.55E-03	4.11E-03	1.19E+01	2.36E+00	1.34E+01	1.34E+01	1.26E+01	1.37E+01	1.36E+01	1.28E+01	1.16E+01	1.20E+01	1.44E+01
	Min	1.11E-03	2.64E-03	1.08E+01	3.47E-07	1.29E+01	1.23E+01	1.13E+01	1.31E+01	1.26E+01	1.21E+01	9.93E+00	8.89E+00	1.40E+01
	Max	2.03E-03	5.39E-03	1.31E+01	1.05E+01	1.37E+01	1.44E+01	1.35E+01	1.41E+01	1.44E+01	1.32E+01	1.30E+01	1.39E+01	1.48E+01
	Std	2.10E-04	6.95E-04	6.14E-01	2.92E+00	1.97E-01	5.05E-01	5.63E-01	2.21E-01	4.33E-01	2.76E-01	7.28E-01	1.30E+00	2.16E-01
17	Mean	1.67E+00	2.40E+00	1.23E+06	5.08E+05	3.01E+07	1.13E+05	2.87E+04	7.71E+07	1.71E+08	2.47E+05	6.70E+05	1.78E+07	4.69E+08
	Min	4.52E-01	1.17E+00	9.23E+04	2.61E+01	1.16E+07	1.53E+04	2.79E+03	2.45E+07	1.63E+07	7.89E+04	1.85E+04	6.17E+01	1.66E+08
	Max	4.07E+00	4.65E+00	4.23E+06	9.25E+06	5.95E+07	4.17E+05	1.20E+05	1.65E+08	5.72E+08	4.95E+05	3.53E+06	1.14E+08	8.79E+08
	Std	9.24E-01	9.35E-01	1.07E+06	1.68E+06	1.09E+07	1.07E+05	2.35E+04	3.83E+07	1.41E+08	1.10E+05	7.90E+05	2.68E+07	1.95E+08
18	Mean	4.70E+01	7.33E+01	4.67E+03	1.29E+07	1.99E+09	5.40E+03	3.25E+03	4.02E+09	4.08E+09	1.56E+04	2.71E+03	5.56E+06	1.09E+10
	Min	2.65E+01	2.42E+01	3.22E+02	2.74E-01	7.14E+08	8.40E+01	5.30E+01	1.54E+09	4.33E+08	1.47E+03	1.45E+02	0.00E+00	4.13E+09
	Max	7.25E+01	1.51E+02	2.29E+04	3.20E+08	3.27E+09	2.63E+04	1.27E+04	6.16E+09	8.35E+09	5.23E+04	1.93E+04	4.26E+07	1.82E+10
	Std	1.31E+01	2.66E+01	4.95E+03	5.82E+07	6.14E+08	7.21E+03	3.68E+03	1.33E+09	2.02E+09	1.37E+04	3.81E+03	1.04E+07	3.57E+09
19	Mean	2.83E-02	4.06E-02	1.66E+01	4.54E+00	2.57E+02	1.46E+02	1.83E+01	4.85E+02	7.63E+02	1.68E+01	3.07E+01	3.59E+01	1.52E+03
	Min	2.03E-02	2.99E-02	1.38E+01	3.94E-03	1.55E+02	7.87E+01	1.45E+01	2.51E+02	2.65E+02	1.38E+01	1.47E+01	0.00E+00	5.50E+02
	Max	3.80E-02	5.10E-02	1.98E+01	2.67E+01	3.32E+02	2.51E+02	2.21E+01	7.23E+02	2.52E+03	2.22E+01	1.12E+02	1.93E+02	3.59E+03
	Std	4.57E-03	5.25E-03	1.85E+00	6.36E+00	4.32E+01	4.01E+01	2.07E+00	1.22E+02	4.72E+02	1.97E+00	2.39E+01	4.14E+01	7.42E+02
20	Mean	5.67E-02	5.20E-02	2.28E+04	2.51E+03	1.55E+05	2.39E+04	1.29E+04	8.23E+05	1.50E+07	4.57E+02	9.83E+03	2.10E+05	8.95E+07
	Min	1.26E-02	4.52E-03	4.01E+03	2.60E+00	4.41E+04	7.30E+03	1.76E+03	1.32E+05	3.82E+04	1.98E+02	1.14E+03	1.78E+02	1.50E+06
	Max	9.79E-02	9.71E-02	9.77E+04	3.78E+04	4.40E+05	4.42E+04	2.86E+04	1.85E+06	1.77E+08	8.80E+02	3.28E+04	2.22E+06	6.64E+08
	Std	2.25E-02	2.30E-02	1.75E+04	7.37E+03	9.78E+04	9.90E+03	7.77E+03	4.90E+05	3.57E+07	1.76E+02	6.52E+03	5.05E+05	1.59E+08
21	Mean	7.45E-01	1.25E+00	4.94E+05	1.71E+05	1.08E+07	7.43E+04	2.93E+04	3.20E+07	9.68E+07	1.87E+04	1.87E+05	3.53E+06	2.65E+08
	Min	1.68E-01	5.45E-01	1.34E+04	1.55E+01	2.02E+06	1.14E+04	2.29E+03	8.00E+06	1.32E+07	5.47E+03	6.75E+03	2.85E+03	5.04E+07
	Max	1.54E+00	2.40E+00	1.81E+06	2.04E+06	1.96E+07	3.79E+05	7.90E+04	7.04E+07	3.31E+08	4.54E+04	1.00E+06	1.29E+07	6.90E+08
	Std	3.01E-01	5.05E-01	4.67E+05	4.24E+05	3.67E+06	8.03E+04	2.35E+04	1.73E+07	8.86E+07	9.66E+03	2.31E+05	4.10E+06	1.83E+08
22	Mean	1.32E-01	1.75E-01	3.13E+02	2.26E+01	1.57E+03	1.67E+03	4.14E+02	2.44E+03	3.69E+03	3.56E+02	6.29E+02	4.75E+02	2.88E+04
	Min	7.97E-02	1.21E-01	3.88E+01	1.50E-01	1.11E+03	8.90E+02	5.08E+01	1.44E+03	7.59E+02	1.11E+02	2.91E+01	1.90E+01	2.73E+03
	Max	1.77E-01	2.10E-01	7.24E+02	1.88E+02	1.94E+03	2.36E+03	1.08E+03	3.86E+03	2.25E+04	5.35E+02	1.04E+03	1.25E+03	1.52E+05
	Std	2.24E-02	2.40E-02	2.08E+02	4.82E+01	2.16E+02	3.09E+02	2.15E+02	5.36E+02	4.39E+03	9.22E+01	2.71E+02	3.96E+02	3.18E+04

23	Mean	2.00E+02	2.00E+02	2.18E+02	2.02E+02	6.52E+02	2.27E+02	2.01E+02	1.02E+03	8.78E+02	2.04E+02	2.26E+02	2.56E+02	1.84E+03
	Min	2.00E+02	2.00E+02	2.00E+02	2.00E+02	5.14E+02	2.09E+02	2.00E+02	5.86E+02	4.31E+02	2.02E+02	2.02E+02	2.03E+02	9.29E+02
	Max	2.00E+02	2.00E+02	2.43E+02	2.09E+02	8.15E+02	2.50E+02	2.11E+02	1.37E+03	2.00E+03	2.07E+02	2.89E+02	3.39E+02	2.64E+03
	Std	8.19E-04	1.78E-03	1.33E+01	2.30E+00	7.41E+01	1.09E+01	2.76E+00	1.71E+02	3.57E+02	1.52E+00	2.20E+01	2.86E+01	4.76E+02
24	Mean	2.00E+02	2.00E+02	2.15E+02	2.01E+02	3.96E+02	3.57E+02	2.25E+02	4.69E+02	4.11E+02	2.24E+02	2.44E+02	2.41E+02	5.78E+02
	Min	2.00E+02	2.00E+02	2.11E+02	2.00E+02	3.44E+02	2.57E+02	2.12E+02	4.19E+02	3.44E+02	2.15E+02	2.22E+02	2.00E+02	4.76E+02
	Max	2.00E+02	2.00E+02	2.25E+02	2.06E+02	4.22E+02	5.74E+02	2.38E+02	5.05E+02	5.11E+02	2.34E+02	2.65E+02	3.10E+02	6.66E+02
	Std	3.85E-02	3.34E-02	3.51E+00	1.27E+00	1.54E+01	8.25E+01	7.21E+00	2.53E+01	3.92E+01	6.40E+00	1.19E+01	3.41E+01	4.67E+01
25	Mean	2.00E+02	2.00E+02	2.06E+02	2.00E+02	2.83E+02	2.36E+02	2.16E+02	3.23E+02	3.54E+02	2.09E+02	2.10E+02	2.30E+02	5.20E+02
	Min	2.00E+02	2.00E+02	2.00E+02	1.99E+02	2.57E+02	2.09E+02	2.06E+02	2.72E+02	2.51E+02	2.04E+02	2.00E+02	2.02E+02	3.32E+02
	Max	2.00E+02	2.00E+02	2.12E+02	2.12E+02	3.10E+02	3.56E+02	2.30E+02	3.65E+02	5.12E+02	2.14E+02	2.23E+02	2.67E+02	8.90E+02
	Std	2.09E-02	2.47E-02	3.03E+00	2.18E+00	1.43E+01	3.09E+01	6.31E+00	2.65E+01	7.23E+01	2.24E+00	5.72E+00	1.64E+01	1.13E+02
26	Mean	1.18E+02	1.86E+02	1.53E+02	1.63E+02	1.08E+02	1.87E+02	1.40E+02	1.45E+02	2.49E+02	1.05E+02	1.88E+02	1.54E+02	4.96E+02
	Min	1.11E+02	1.07E+02	1.05E+02	1.10E+02	1.08E+02	1.06E+02	1.05E+02	1.11E+02	1.12E+02	1.05E+02	1.05E+02	1.08E+02	1.77E+02
	Max	1.36E+02	2.00E+02	2.02E+02	2.14E+02	1.09E+02	4.45E+02	2.00E+02	3.40E+02	4.33E+02	1.05E+02	2.02E+02	4.06E+02	9.16E+02
	Std	7.31E+00	3.02E+01	4.89E+01	4.44E+01	2.67E-01	6.31E+01	4.64E+01	4.68E+01	1.21E+02	2.60E-02	3.25E+01	8.35E+01	1.50E+02
27	Mean	2.00E+02	1.10E+03	5.45E+02	3.90E+02	1.36E+03	1.73E+03	5.84E+02	1.40E+03	1.56E+03	5.60E+02	9.11E+02	1.22E+03	1.87E+03
	Min	2.00E+02	6.56E+02	4.47E+02	2.00E+02	1.22E+03	7.06E+02	4.46E+02	1.03E+03	1.19E+03	4.80E+02	4.61E+02	5.58E+02	1.49E+03
	Max	2.00E+02	1.40E+03	6.98E+02	1.31E+03	1.43E+03	2.31E+03	7.75E+02	1.63E+03	1.73E+03	7.72E+02	1.24E+03	1.68E+03	2.45E+03
	Std	3.74E-03	2.33E+02	7.00E+01	3.18E+02	5.97E+01	3.03E+02	8.58E+01	1.51E+02	1.07E+02	7.83E+01	3.12E+02	2.69E+02	2.19E+02
28	Mean	2.00E+02	4.27E+03	1.00E+03	1.22E+03	2.17E+03	4.74E+03	2.19E+03	6.58E+03	6.78E+03	1.11E+03	2.39E+03	3.99E+03	9.10E+03
	Min	2.00E+02	2.00E+02	7.16E+02	2.00E+02	1.88E+03	3.16E+03	8.63E+02	5.08E+03	4.12E+03	8.95E+02	1.13E+03	2.14E+03	6.95E+03
	Max	2.00E+02	6.96E+03	2.14E+03	6.26E+03	2.47E+03	6.47E+03	5.29E+03	7.70E+03	9.10E+03	1.84E+03	4.70E+03	8.07E+03	1.18E+04
	Std	5.12E-03	1.29E+03	3.84E+02	1.72E+03	1.60E+02	8.21E+02	1.38E+03	6.55E+02	1.35E+03	2.13E+02	7.63E+02	1.42E+03	1.27E+03
29	Mean	2.00E+02	2.00E+02	1.17E+04	8.79E+04	1.49E+07	1.72E+07	8.32E+02	1.66E+08	3.99E+05	1.80E+04	4.60E+05	9.71E+07	5.33E+08
	Min	2.00E+02	2.00E+02	4.73E+03	2.00E+02	8.32E+06	1.09E+03	3.73E+02	6.45E+07	2.63E+04	2.23E+03	3.74E+02	2.99E+04	2.38E+08
	Max	2.00E+02	2.00E+02	2.23E+04	1.51E+06	2.08E+07	2.38E+07	2.20E+03	2.66E+08	1.35E+06	3.58E+05	5.54E+06	3.33E+08	9.70E+08
	Std	3.94E-03	9.17E-03	4.06E+03	2.88E+05	3.11E+06	9.85E+06	3.79E+02	5.58E+07	3.02E+05	6.43E+04	1.34E+06	7.25E+07	2.02E+08
30	Mean	2.00E+02	2.00E+02	1.17E+04	6.56E+02	5.00E+05	7.00E+03	2.62E+03	2.19E+06	3.76E+06	3.43E+03	4.06E+03	4.47E+05	9.55E+06
	Min	2.00E+02	2.00E+02	3.53E+03	2.00E+02	1.94E+05	2.11E+03	1.99E+03	1.07E+06	8.44E+04	2.12E+03	1.66E+03	3.79E+03	2.91E+06
	Max	2.00E+02	2.00E+02	3.27E+04	3.74E+03	9.55E+05	6.18E+04	3.77E+03	3.52E+06	1.12E+07	5.82E+03	8.37E+03	3.99E+06	2.19E+07
	Std	2.01E-02	1.97E-02	7.12E+03	8.31E+02	2.04E+05	1.15E+04	5.01E+02	6.04E+05	3.08E+06	7.87E+02	1.78E+03	8.86E+05	4.96E+06

Appendix IV: CEC Test suite 2014 results for 50 dimensions

		SRFA	RFA	FA	ODFA	CFA	LSFA	SFA	SA	BSO	CSO	PSO	DFO	ALO
1	Mean	3.71E+01	1.20E+02	1.83E+07	2.89E+04	1.33E+09	2.86E+07	5.01E+05	3.22E+09	1.70E+09	4.95E+07	8.65E+06	1.78E+08	5.58E+09
	Min	1.53E+01	5.65E+01	6.02E+06	1.54E-03	4.36E+08	2.80E+06	4.81E+04	1.49E+09	8.21E+08	3.38E+07	2.27E+06	8.44E+06	3.32E+09
	Max	6.32E+01	1.77E+02	3.21E+07	2.74E+05	1.88E+09	2.17E+08	2.95E+06	4.51E+09	3.68E+09	6.76E+07	2.51E+07	5.91E+08	8.90E+09
	Std	9.75E+00	2.72E+01	6.05E+06	6.45E+04	3.06E+08	4.60E+07	5.84E+05	7.48E+08	7.47E+08	8.91E+06	5.63E+06	1.39E+08	1.29E+09
2	Mean	1.45E+03	5.84E+03	1.02E+05	8.43E+05	6.29E+10	6.88E+09	2.47E+03	1.59E+11	5.39E+10	1.00E+10	6.49E+07	4.11E+09	1.89E+11
	Min	1.09E+03	4.36E+03	5.14E+04	4.19E+00	4.46E+10	2.99E+09	9.26E+00	1.31E+11	2.94E+10	1.00E+10	4.35E+05	2.66E+08	1.47E+11
	Max	1.81E+03	6.93E+03	1.48E+05	1.21E+07	7.36E+10	1.26E+10	1.30E+04	1.78E+11	1.14E+11	1.00E+10	5.69E+08	1.16E+10	2.26E+11
	Std	1.73E+02	5.73E+02	2.01E+04	2.26E+06	6.77E+09	2.06E+09	2.85E+03	1.11E+10	2.06E+10	0.00E+00	1.23E+08	3.31E+09	1.79E+10
3	Mean	4.89E-03	1.27E-02	6.63E+04	8.02E+01	1.24E+05	2.95E+04	8.64E+03	2.87E+05	4.00E+05	5.41E+04	2.14E+04	1.24E+05	4.17E+06
	Min	2.20E-03	7.13E-03	4.46E+04	1.41E-04	8.22E+04	1.63E+04	3.00E+03	1.91E+05	9.76E+04	4.04E+04	4.89E+03	4.19E+04	2.58E+05
	Max	8.87E-03	2.29E-02	8.41E+04	6.01E+02	1.51E+05	5.43E+04	1.54E+04	7.21E+05	1.68E+06	8.28E+04	4.80E+04	3.19E+05	2.57E+07
	Std	1.80E-03	2.98E-03	1.05E+04	1.56E+02	1.87E+04	8.31E+03	3.78E+03	1.22E+05	3.29E+05	9.54E+03	1.11E+04	5.95E+04	6.97E+06
4	Mean	3.06E-04	1.11E-03	5.00E+01	1.09E-01	9.81E+03	8.36E+02	5.04E+01	3.79E+04	1.01E+04	2.82E+02	1.53E+02	8.23E+02	5.96E+04
	Min	2.11E-04	8.14E-04	3.99E+01	6.69E-08	5.53E+03	3.52E+02	3.85E+01	2.73E+04	4.74E+03	1.56E+02	4.79E+01	1.57E+00	4.16E+04
	Max	3.65E-04	1.52E-03	9.97E+01	1.32E+00	1.32E+04	1.78E+03	1.07E+02	5.17E+04	3.20E+04	3.56E+02	3.40E+02	3.69E+03	8.17E+04
	Std	3.75E-05	1.55E-04	1.13E+01	3.21E-01	1.69E+03	3.15E+02	1.81E+01	6.49E+03	5.24E+03	3.97E+01	5.55E+01	7.60E+02	1.12E+04
5	Mean	2.35E-02	4.88E-02	2.01E+01	7.10E-01	2.12E+01	2.00E+01	2.00E+01	2.13E+01	2.00E+01	2.12E+01	2.06E+01	2.11E+01	2.15E+01
	Min	2.00E-02	4.43E-02	2.01E+01	1.94E-04	2.11E+01	2.00E+01	2.00E+01	2.12E+01	2.00E+01	2.11E+01	2.00E+01	2.07E+01	2.14E+01
	Max	2.66E-02	5.34E-02	2.03E+01	5.04E+00	2.13E+01	2.00E+01	2.00E+01	2.14E+01	2.00E+01	2.13E+01	2.12E+01	2.14E+01	2.16E+01
	Std	1.28E-03	2.17E-03	5.76E-02	1.14E+00	5.38E-02	3.98E-05	3.34E-05	5.53E-02	3.79E-05	3.28E-02	4.35E-01	1.88E-01	5.96E-02
6	Mean	3.16E-01	4.83E-01	3.65E+00	1.36E+00	6.89E+01	8.13E+01	4.36E+00	7.97E+01	7.13E+01	5.85E+01	4.11E+01	4.64E+01	8.54E+01
	Min	2.93E-01	4.57E-01	1.34E+00	3.61E-02	6.49E+01	7.36E+01	1.47E+00	7.59E+01	5.38E+01	5.37E+01	2.87E+01	3.17E+00	8.17E+01
	Max	3.33E-01	5.13E-01	7.43E+00	4.12E+00	7.13E+01	8.56E+01	8.27E+00	8.35E+01	7.96E+01	6.29E+01	5.03E+01	7.83E+01	9.06E+01
	Std	1.01E-02	1.34E-02	1.65E+00	1.01E+00	1.66E+00	3.02E+00	1.83E+00	1.78E+00	5.61E+00	1.92E+00	4.22E+00	2.00E+01	1.90E+00
7	Mean	1.82E-03	7.10E-03	2.47E-01	6.99E-01	6.04E+02	1.55E+02	1.49E-01	1.51E+03	4.82E+02	4.23E+00	2.94E+00	3.01E+01	1.86E+03
	Min	1.20E-03	5.50E-03	1.57E-01	1.35E-08	5.21E+02	7.49E+01	4.31E-04	1.36E+03	2.46E+02	2.73E+00	1.22E+00	7.18E-02	1.63E+03
	Max	2.34E-03	9.27E-03	3.76E-01	9.67E+00	7.42E+02	2.21E+02	4.64E-01	1.67E+03	9.06E+02	8.89E+00	1.01E+01	1.19E+02	2.13E+03
	Std	2.85E-04	1.05E-03	5.32E-02	1.75E+00	4.84E+01	3.53E+01	1.23E-01	8.61E+01	1.61E+02	1.12E+00	1.88E+00	3.02E+01	1.14E+02
8	Mean	4.95E-04	1.96E-03	5.44E+01	5.50E-01	5.20E+02	3.49E+02	2.11E+02	7.04E+02	4.91E+02	2.28E+02	1.05E+02	2.84E+02	7.75E+02
	Min	4.14E-04	1.39E-03	3.39E+01	1.79E-05	4.76E+02	2.66E+02	1.32E+02	6.23E+02	4.05E+02	1.82E+02	6.93E+01	7.79E+01	7.12E+02
	Max	5.92E-04	2.27E-03	7.46E+01	8.84E+00	5.59E+02	4.23E+02	2.78E+02	7.44E+02	5.95E+02	2.83E+02	1.47E+02	4.86E+02	8.35E+02
	Std	4.80E-05	2.01E-04	1.10E+01	1.71E+00	1.86E+01	4.25E+01	4.17E+01	2.80E+01	4.62E+01	2.14E+01	2.03E+01	1.07E+02	3.45E+01
9	Mean	8.29E-04	3.30E-03	5.83E+01	2.86E-01	6.21E+02	5.54E+02	2.77E+02	8.95E+02	5.34E+02	2.90E+02	1.42E+02	2.36E+02	1.02E+03
	Min	6.33E-04	2.49E-03	2.79E+01	1.58E-09	5.64E+02	4.56E+02	1.94E+02	8.15E+02	4.74E+02	2.42E+02	1.09E+02	0.00E+00	9.13E+02
	Max	1.03E-03	3.85E-03	1.19E+02	2.41E+00	6.65E+02	6.57E+02	3.87E+02	9.71E+02	7.29E+02	3.50E+02	1.96E+02	3.90E+02	1.14E+03
	Std	1.04E-04	3.24E-04	1.95E+01	5.69E-01	2.26E+01	4.85E+01	4.84E+01	4.27E+01	5.56E+01	2.48E+01	2.60E+01	1.09E+02	5.73E+01
10	Mean	1.19E-02	4.53E-02	3.95E+03	2.19E+00	1.41E+04	8.49E+03	7.25E+03	1.59E+04	1.44E+04	7.97E+03	5.40E+03	4.15E+03	1.75E+04
	Min	8.35E-03	3.24E-02	2.12E+03	1.70E-07	1.32E+04	6.67E+03	5.67E+03	1.52E+04	1.17E+04	7.32E+03	3.13E+03	2.00E-11	1.56E+04
	Max	1.49E-02	5.32E-02	6.72E+03	1.79E+01	1.49E+04	9.93E+03	8.74E+03	1.67E+04	1.62E+04	8.74E+03	7.02E+03	1.05E+04	1.87E+04
	Std	1.54E-03	5.10E-03	1.07E+03	3.84E+00	4.34E+02	8.18E+02	8.05E+02	3.81E+02	1.05E+03	3.43E+02	9.16E+02	3.09E+03	6.40E+02
11	Mean	2.05E-02	7.78E-02	4.11E+03	3.25E+01	1.42E+04	7.78E+03	7.43E+03	1.55E+04	1.32E+04	8.96E+03	6.56E+03	1.16E+04	1.70E+04
	Min	1.43E-02	5.70E-02	2.51E+03	1.55E-07	1.32E+04	5.64E+03	5.25E+03	1.42E+04	8.69E+03	8.14E+03	4.12E+03	5.04E+03	1.59E+04

	Max	2.56E-02	9.30E-02	6.43E+03	9.02E+02	1.50E+04	9.88E+03	9.66E+03	1.64E+04	1.61E+04	9.79E+03	8.66E+03	1.51E+04	1.79E+04
	Std	2.30E-03	9.20E-03	9.11E+02	1.64E+02	4.24E+02	9.37E+02	1.12E+03	5.14E+02	1.82E+03	3.20E+02	1.04E+03	2.36E+03	5.66E+02
12	Mean	1.08E-03	1.46E-01	5.29E-02	2.80E-02	4.21E+00	8.70E-01	7.60E-01	5.77E+00	4.02E+00	2.20E+00	1.57E+00	3.00E+00	7.50E+00
	Min	9.44E-04	1.71E-03	2.56E-02	5.35E-06	3.27E+00	3.15E-01	3.26E-01	4.30E+00	3.11E+00	1.85E+00	8.32E-01	2.39E-01	4.88E+00
	Max	1.20E-03	4.31E+00	1.82E-01	1.89E-01	4.73E+00	1.52E+00	1.15E+00	7.57E+00	5.65E+00	2.63E+00	2.44E+00	5.24E+00	9.28E+00
	Std	6.08E-05	7.86E-01	2.95E-02	4.13E-02	3.41E-01	3.01E-01	2.51E-01	8.24E-01	6.49E-01	1.97E-01	3.86E-01	1.27E+00	9.66E-01
13	Mean	5.01E-02	4.38E-02	3.92E-01	2.05E-01	5.13E+00	4.64E+00	5.82E-01	8.40E+00	4.53E+00	3.46E-01	5.63E-01	4.92E-01	9.81E+00
	Min	1.59E-02	1.30E-02	2.77E-01	5.55E-02	4.43E+00	3.99E+00	3.58E-01	7.54E+00	3.44E+00	2.57E-01	3.89E-01	0.00E+00	9.03E+00
	Max	7.52E-02	6.26E-02	5.57E-01	3.34E-01	5.60E+00	5.15E+00	8.65E-01	9.22E+00	6.60E+00	4.42E-01	8.08E-01	2.84E+00	1.07E+01
	Std	1.30E-02	1.10E-02	6.93E-02	7.84E-02	2.64E-01	3.23E-01	1.11E-01	4.38E-01	8.11E-01	4.18E-02	9.53E-02	7.50E-01	4.57E-01
14	Mean	1.47E-02	1.39E-02	4.68E-01	9.41E-02	1.82E+02	6.90E+01	5.04E-01	4.32E+02	1.52E+02	2.93E-01	3.35E-01	1.72E+01	5.26E+02
	Min	1.07E-03	4.26E-03	3.35E-01	7.34E-03	1.40E+02	3.41E+01	3.00E-01	3.58E+02	8.69E+01	2.10E-01	2.37E-01	0.00E+00	4.26E+02
	Max	3.79E-02	4.24E-02	8.08E-01	2.32E-01	2.08E+02	9.98E+01	8.37E-01	4.75E+02	2.05E+02	3.47E-01	9.32E-01	5.52E+01	6.29E+02
	Std	8.80E-03	8.09E-03	1.45E-01	5.73E-02	1.82E+01	1.53E+01	2.01E-01	3.11E+01	2.73E+01	2.81E-02	1.62E-01	1.73E+01	5.61E+01
15	Mean	8.26E-08	1.19E-06	7.80E+00	1.61E+00	7.92E+05	8.93E+03	2.47E+01	1.87E+07	1.12E+06	5.05E+01	1.45E+02	5.04E+03	4.56E+07
	Min	5.35E-08	7.20E-07	5.42E+00	0.00E+00	3.03E+05	6.07E+03	9.19E+00	3.83E+06	3.41E+04	3.81E+01	5.13E+01	6.36E+01	4.16E+06
	Max	1.37E-07	1.64E-06	1.06E+01	3.34E+01	1.58E+06	1.08E+04	3.98E+01	3.07E+07	1.08E+07	7.26E+01	4.30E+02	1.15E+05	7.74E+07
	Std	1.68E-08	2.36E-07	1.27E+00	6.32E+00	3.06E+05	9.92E+02	1.03E+01	6.08E+06	2.21E+06	7.02E+00	7.81E+01	2.09E+04	1.78E+07
16	Mean	3.16E-03	1.15E-02	1.99E+01	1.51E+00	2.27E+01	2.27E+01	2.14E+01	2.34E+01	2.26E+01	2.24E+01	1.96E+01	2.10E+01	2.41E+01
	Min	2.55E-03	7.93E-03	1.73E+01	8.33E-08	2.22E+01	2.11E+01	1.92E+01	2.28E+01	2.00E+01	2.17E+01	1.74E+01	4.24E+00	2.36E+01
	Max	4.04E-03	1.40E-02	2.16E+01	1.73E+01	2.32E+01	2.38E+01	2.31E+01	2.44E+01	2.41E+01	2.29E+01	2.15E+01	2.31E+01	2.44E+01
	Std	3.55E-04	1.45E-03	1.04E+00	3.57E+00	1.98E-01	5.83E-01	9.39E-01	3.25E-01	1.01E+00	2.68E-01	1.26E+00	3.36E+00	2.39E-01
17	Mean	4.05E+00	9.72E+00	2.80E+06	1.21E+05	1.07E+08	1.35E+06	9.18E+04	2.94E+08	2.45E+08	3.63E+06	1.25E+06	2.01E+07	8.34E+08
	Min	2.17E+00	3.00E+00	8.79E+05	2.74E-03	5.58E+07	1.20E+05	2.17E+04	7.79E+07	7.57E+07	2.10E+06	1.41E+05	8.21E+04	2.78E+08
	Max	7.27E+00	1.94E+01	7.91E+06	1.36E+06	1.71E+08	5.88E+06	2.84E+05	5.51E+08	8.14E+08	5.33E+06	7.23E+06	1.60E+08	1.91E+09
	Std	1.16E+00	3.24E+00	2.03E+06	3.07E+05	2.98E+07	1.21E+06	7.09E+04	1.07E+08	1.91E+08	9.44E+05	1.42E+06	3.01E+07	3.80E+08
18	Mean	1.36E+02	3.33E+02	3.23E+03	1.37E+05	5.08E+09	2.24E+08	3.81E+03	1.49E+10	6.76E+09	5.58E+09	1.56E+03	7.65E+07	2.40E+10
	Min	6.85E+01	1.81E+02	5.21E+02	3.49E-03	2.74E+09	8.08E+03	2.22E+02	6.80E+09	1.87E+09	1.93E+08	2.28E+02	0.00E+00	7.40E+09
	Max	1.90E+02	4.94E+02	1.56E+04	3.02E+06	6.70E+09	8.50E+08	1.24E+04	1.91E+10	1.35E+10	1.00E+10	7.61E+03	9.52E+08	3.06E+10
	Std	2.78E+01	9.19E+01	3.39E+03	5.60E+05	1.07E+09	2.45E+08	3.35E+03	2.81E+09	2.49E+09	4.00E+09	1.79E+03	1.89E+08	4.71E+09
19	Mean	5.45E-02	8.49E-02	2.77E+01	1.61E+00	4.94E+02	1.91E+02	3.09E+01	1.93E+03	1.08E+03	3.88E+01	5.62E+01	1.15E+02	3.50E+03
	Min	3.99E-02	6.99E-02	2.32E+01	1.18E-03	2.97E+02	1.29E+02	2.67E+01	9.08E+02	3.22E+02	3.35E+01	2.11E+01	6.07E+00	2.19E+03
	Max	6.21E-02	9.77E-02	3.32E+01	8.22E+00	6.63E+02	3.38E+02	3.71E+01	2.87E+03	3.98E+03	4.40E+01	1.32E+02	3.96E+02	6.24E+03
	Std	4.85E-03	8.19E-03	2.51E+00	2.14E+00	9.35E+01	4.56E+01	2.45E+00	4.62E+02	6.34E+02	2.69E+00	3.33E+01	1.07E+02	1.05E+03
20	Mean	5.87E-02	7.08E-02	3.10E+04	1.53E+03	1.97E+05	4.57E+04	1.34E+04	1.88E+06	2.75E+06	2.16E+04	2.30E+04	1.01E+05	3.35E+07
	Min	2.36E-02	3.00E-02	9.03E+03	3.55E-01	4.31E+04	1.87E+04	3.91E+03	2.76E+05	1.17E+05	8.93E+03	6.58E+03	2.42E+01	6.14E+05
	Max	1.15E-01	1.30E-01	6.27E+04	1.63E+04	4.12E+05	8.62E+04	2.61E+04	6.34E+06	1.59E+07	4.74E+04	5.26E+04	5.99E+05	1.63E+08
	Std	2.32E-02	2.72E-02	1.27E+04	3.06E+03	1.07E+05	1.80E+04	6.16E+03	1.49E+06	3.53E+06	9.30E+03	1.30E+04	1.26E+05	4.42E+07
21	Mean	2.14E+00	4.77E+00	9.38E+05	1.16E+04	4.50E+07	1.31E+06	5.07E+04	1.29E+08	1.84E+08	9.20E+05	4.81E+05	1.11E+07	5.17E+08
	Min	9.55E-01	1.31E+00	1.45E+05	1.26E+01	1.96E+07	9.31E+04	7.21E+03	5.60E+07	1.78E+07	4.27E+05	4.05E+04	3.35E+03	2.28E+08
	Max	4.12E+00	9.36E+00	2.78E+06	1.41E+05	8.09E+07	1.24E+07	2.69E+05	2.10E+08	4.85E+08	1.49E+06	2.14E+06	5.41E+07	1.25E+09
	Std	7.19E-01	1.82E+00	6.91E+05	2.65E+04	1.31E+07	2.22E+06	5.04E+04	4.68E+07	1.20E+08	2.69E+05	5.59E+05	1.41E+07	2.33E+08
22	Mean	1.11E-01	1.68E-01	5.02E+02	3.26E+00	3.29E+03	3.38E+03	9.07E+02	1.50E+04	4.33E+03	1.02E+03	8.83E+02	1.04E+03	7.83E+04
	Min	6.08E-02	1.20E-01	1.55E+02	7.98E-02	2.60E+03	2.39E+03	9.78E+01	5.03E+03	2.19E+03	6.74E+02	3.02E+02	4.95E+00	5.41E+03
	Max	1.42E-01	2.07E-01	1.24E+03	1.94E+01	3.93E+03	4.82E+03	1.60E+03	3.49E+04	1.01E+04	1.35E+03	1.54E+03	2.37E+03	2.01E+05
	Std	1.84E-02	1.93E-02	2.98E+02	4.98E+00	3.62E+02	6.28E+02	3.30E+02	7.46E+03	2.27E+03	1.63E+02	3.47E+02	6.54E+02	5.44E+04

X

23	Mean	2.00E+02	2.00E+02	2.25E+02	2.00E+02	7.43E+02	3.45E+02	2.07E+02	1.86E+03	8.85E+02	2.16E+02	2.19E+02	2.53E+02	2.97E+03
	Min	2.00E+02	2.00E+02	2.04E+02	2.00E+02	5.12E+02	2.46E+02	2.00E+02	1.33E+03	5.20E+02	2.12E+02	2.01E+02	2.00E+02	1.49E+03
	Max	2.00E+02	2.00E+02	2.60E+02	2.01E+02	8.75E+02	4.41E+02	2.28E+02	2.35E+03	1.54E+03	2.22E+02	2.51E+02	3.81E+02	4.19E+03
	Std	6.19E-04	1.17E-03	1.13E+01	2.91E-01	8.70E+01	4.66E+01	7.31E+00	2.80E+02	2.31E+02	2.75E+00	1.62E+01	4.39E+01	7.33E+02
24	Mean	2.00E+02	2.00E+02	2.53E+02	2.00E+02	4.97E+02	5.40E+02	2.59E+02	7.92E+02	5.95E+02	2.81E+02	2.95E+02	2.99E+02	9.13E+02
	Min	2.00E+02	2.00E+02	2.49E+02	2.00E+02	4.44E+02	3.93E+02	2.52E+02	7.35E+02	4.41E+02	2.74E+02	2.74E+02	2.00E+02	7.01E+02
	Max	2.00E+02	2.00E+02	2.60E+02	2.01E+02	5.32E+02	7.67E+02	2.69E+02	8.75E+02	8.52E+02	2.88E+02	3.45E+02	4.05E+02	1.04E+03
	Std	4.07E-02	3.78E-02	2.29E+00	3.18E-01	1.98E+01	8.86E+01	4.61E+00	3.48E+01	9.83E+01	3.68E+00	1.70E+01	6.10E+01	7.43E+01
25	Mean	2.00E+02	2.00E+02	2.10E+02	2.00E+02	3.71E+02	2.63E+02	2.20E+02	5.63E+02	4.41E+02	2.26E+02	2.22E+02	2.33E+02	8.17E+02
	Min	2.00E+02	2.00E+02	2.04E+02	2.00E+02	3.15E+02	2.31E+02	2.07E+02	4.45E+02	2.96E+02	2.21E+02	2.11E+02	2.01E+02	5.68E+02
	Max	2.00E+02	2.00E+02	2.17E+02	2.00E+02	4.29E+02	3.57E+02	2.33E+02	7.24E+02	8.12E+02	2.35E+02	2.38E+02	3.00E+02	1.10E+03
	Std	1.10E-02	1.34E-02	2.98E+00	8.65E-02	2.44E+01	2.69E+01	6.77E+00	7.27E+01	1.19E+02	3.63E+00	7.78E+00	2.23E+01	1.52E+02
26	Mean	1.84E+02	1.97E+02	1.93E+02	1.68E+02	1.12E+02	1.92E+02	1.88E+02	4.84E+02	2.73E+02	1.07E+02	1.89E+02	3.11E+02	8.85E+02
	Min	1.22E+02	1.16E+02	1.06E+02	1.14E+02	1.10E+02	1.10E+02	1.06E+02	1.31E+02	1.36E+02	1.07E+02	1.09E+02	1.17E+02	5.28E+02
	Max	2.00E+02	2.00E+02	2.06E+02	2.00E+02	1.13E+02	2.67E+02	2.01E+02	7.65E+02	5.30E+02	1.07E+02	2.03E+02	6.61E+02	1.62E+03
	Std	2.54E+01	1.53E+01	2.94E+01	3.64E+01	4.81E-01	4.73E+01	3.26E+01	1.55E+02	1.15E+02	2.12E-01	3.13E+01	1.82E+02	2.10E+02
27	Mean	2.00E+02	1.67E+03	7.34E+02	2.71E+02	2.25E+03	3.12E+03	7.92E+02	2.66E+03	2.46E+03	1.56E+03	1.85E+03	2.21E+03	3.19E+03
	Min	2.00E+02	1.49E+03	5.47E+02	2.00E+02	2.15E+03	2.76E+03	5.96E+02	2.40E+03	2.32E+03	9.99E+02	1.64E+03	1.96E+03	2.58E+03
	Max	2.00E+02	1.91E+03	9.29E+02	1.29E+03	2.33E+03	3.93E+03	1.05E+03	2.87E+03	2.59E+03	1.97E+03	2.12E+03	2.49E+03	4.23E+03
	Std	4.54E-03	1.25E+02	8.82E+01	2.33E+02	4.46E+01	2.95E+02	1.05E+02	1.15E+02	7.12E+01	3.12E+02	1.07E+02	1.34E+02	4.14E+02
28	Mean	2.00E+02	6.02E+03	1.32E+03	2.32E+02	6.08E+03	9.03E+03	3.64E+03	1.38E+04	1.25E+04	1.60E+03	4.26E+03	8.44E+03	1.66E+04
	Min	2.00E+02	2.99E+03	1.06E+03	2.00E+02	4.64E+03	6.40E+03	1.44E+03	1.19E+04	7.05E+03	1.28E+03	2.30E+03	4.51E+03	1.39E+04
	Max	2.00E+02	1.04E+04	2.87E+03	4.90E+02	7.46E+03	1.17E+04	9.63E+03	1.51E+04	1.76E+04	2.44E+03	6.17E+03	1.40E+04	1.96E+04
	Std	5.67E-03	1.98E+03	3.38E+02	7.01E+01	6.44E+02	1.41E+03	2.69E+03	8.79E+02	2.38E+03	3.05E+02	1.17E+03	2.23E+03	1.54E+03
29	Mean	2.00E+02	2.00E+02	3.49E+04	1.54E+03	1.34E+08	4.61E+08	4.58E+03	8.08E+08	3.42E+06	1.11E+06	4.17E+06	2.88E+08	1.56E+09
	Min	2.00E+02	2.00E+02	1.25E+04	2.00E+02	8.40E+07	4.57E+08	1.01E+03	5.23E+08	6.88E+05	3.43E+05	1.12E+03	1.36E+07	6.41E+08
	Max	2.00E+02	2.00E+02	8.27E+04	2.02E+04	1.78E+08	4.65E+08	8.22E+03	1.13E+09	1.55E+07	4.13E+06	6.17E+07	5.52E+08	2.22E+09
	Std	4.73E-03	1.81E-02	1.49E+04	3.87E+03	2.26E+07	2.17E+06	1.75E+03	1.46E+08	3.08E+06	8.32E+05	1.25E+07	1.31E+08	3.54E+08
30	Mean	2.00E+02	2.00E+02	1.78E+04	2.05E+02	2.19E+06	3.19E+04	3.57E+03	9.71E+06	8.80E+06	1.88E+04	1.23E+04	3.30E+05	2.49E+07
	Min	2.00E+02	2.00E+02	6.61E+03	2.00E+02	6.71E+05	4.80E+03	2.07E+03	3.51E+06	1.04E+06	1.03E+04	3.27E+03	1.20E+03	1.34E+07
	Max	2.00E+02	2.00E+02	4.09E+04	2.90E+02	3.66E+06	9.62E+04	6.00E+03	1.56E+07	2.25E+07	3.42E+04	8.94E+04	1.61E+06	4.53E+07
	Std	1.57E-02	1.58E-02	9.15E+03	1.62E+01	7.98E+05	2.58E+04	9.33E+02	3.09E+06	4.70E+06	5.27E+03	2.08E+04	3.96E+05	8.28E+06